

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:49:52 ; Search time 17.5 Seconds
(without alignments)
164.861 Million cell updates/sec

Title: FVIII_ARG355I

Perfect score: 161

Sequence: 1 AVVKVDSCEPPEPQLIMKNEAEYDDDLT 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154	95.7	2351	1 EZHU	coagulation factor
2	113.5	70.5	2319	2 A47004	coagulation factor
3	90.5	56.2	2133	2 T42763	coagulation factor
4	56	34.8	1102	2 S65235	probable membrane
5	53.5	33.2	390	2 T46028	hypothetical prote
6	53	32.9	321	2 H71924	hypothetical prote
7	53	32.9	6642	2 T29757	protein UNC-89 - C
8	52	32.3	537	2 T31178	hypothetical prote
9	52	32.3	649	2 I64847	epithelial sodium
10	52	32.3	649	2 T38204	epithelial anilori
11	51	31.7	162	2 A55969	FP21 protein - sli
12	51	31.7	428	2 C84476	hypothetical prote
13	51	31.7	973	2 A97522	ribonuclease E, RN
14	51	31.7	977	2 AC2741	ribonuclease E, RN
15	50.5	31.4	141	2 T46015	hypothetical prote
16	50	31.1	161	2 T45459	hypothetical prote
17	50	31.1	333	2 AE0222	skpl homolog - fis
18	50	31.1	475	2 F64151	flagellar motor sw
19	50	31.1	1170	1 TSHUP1	thrombospondin 1 p
20	50	31.1	1170	2 A40558	thrombospondin 1 p
21	49.5	30.7	441	2 G82612	hypothetical prote
22	49.5	30.7	736	2 G72621	probable translati
23	49	30.4	155	2 T03087	conserved hypothet
24	48.5	30.1	521	2 A53153	glucose transport
25	48	29.8	165	2 G88987	protein C50H11.17
26	48	29.8	324	2 D84707	hypothetical prote
27	48	29.8	338	2 T18715	hypothetical prote
28	48	29.8	465	2 S69038	hypothetical prote
29	48	29.8	556	2 T23172	hypothetical prote

ALIGNMENTS

RESULT 1

EZHU

coagulation factor VIII precursor [validated] - human
N: Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant C
C: Species: Homo sapiens (man)
C: Date: 28-Aug-1985 #sequence revision 28-Aug-1985 #text_change 08-Dec-2000
C: Accession: I54318; A00525; I58059; A23584; A26174; A43348; A43986; S63527; S66445;
R: Gitschler, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A: Title: Sequence of the exon-containing regions of the human factor VIII gene.
A: Reference number: I54318; MUID: 93265012; PMID: 1303178
A: Accession: I54318
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1921, 'S', 1923-2351 <RES>
A: Cross-references: GB:M88648; NID: g182381; PIDN: AAA52420.1; PID: g182383
R: Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seel
Nature 312, 330-337, 1984
A: Title: Expression of active human factor VIII from recombinant DNA clones.
A: Reference number: A00525; MUID: 85061548; PMID: 6438526
A: Accession: A00525
A: Molecule type: mRNA
A: Residues: 1-2351 <WOL>
A: Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R: Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.I.
S, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A: Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A: Reference number: I58059; MUID: 85061550; PMID: 6438528
A: Accession: I58059
A: Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
A: Cross-references: GB:K01740; NID: g182802; PIDN: AAA52484.1; PID: g182803
R: Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; J.
B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favallaro, J.; Hansen, J.
DNA 4, 333-349, 1985
A: Title: Characterization of the polypeptide composition of human factor VIII: C and
A: Reference number: A23584; MUID: 86081164; PMID: 3935400
A: Accession: A23584
A: Molecule type: mRNA
A: Residues: 1-2351 <TRU>
A: Cross-references: GB:M14113; NID: g182817; PIDN: AAA52485.1; PID: g182818
R: Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A: Title: Proteolytic processing of human factor VIII. Correlation of specific cleava
ity.
A: Reference number: A26174; MUID: 86159740; PMID: 3082357
A: Accession: A26174
A: Molecule type: protein
A: Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <E
R: Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

30	48	29.8	567	2	T33400	protein kinase C h
31	48	29.8	597	2	T33399	protein kinase C h
32	48	29.8	682	2	T23813	hypothetical prote
33	48	29.8	699	2	T18426	hypothetical prote
34	48	29.8	704	1	S60117	protein kinase C (
35	48	29.8	734	1	S25166	translation elonga
36	48	29.8	746	2	S61057	GYP7 protein - yea
37	48	29.8	1038	2	T23046	hypothetical prote
38	48	29.8	1173	2	T51892	hypothetical prote
39	47.5	29.5	167	2	S50808	conserved hypothet
40	47.5	29.5	830	2	C69011	origin recognition
41	47.5	29.5	899	2	C71608	hypothetical prote
42	47.5	29.5	1046	2	T29776	hypothetical prote
43	47	29.2	145	2	A11572	Salmonella enteric
44	47	29.2	204	2	T23362	hypothetical prote
45	47	29.2	305	2	F86744	tagatose-6-phospha

Biochemistry 31, 3315-3325, 1992
 A:Title: Identification and functional importance of tyrosine sulfate residues within re
 A:Reference number: A42348; MUID:92207952; PMID:1554716
 A:Accession: A42348
 A:Molecule type: protein
 A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
 A:Experimental source: recombinant material from Chinese hamster ovary cells
 A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 J; Pay, P. J.; Smudzin, T. M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A:Reference number: A43986; MUID:89340500; PMID:2503509
 A:Accession: A43986
 A:Molecule type: protein
 A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <P>
 R;Leyte, A.; van Schijndel, H. B.; Niehrs, C.; Huttner, W. B.; Verbeet, M. P.; Mertens, K.;
 J. Biol. Chem. 266, 740-746, 1991
 A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A:Reference number: A56109; MUID:91093266; PMID:1898735
 A:Contents: annotation; sulfation
 R;Gitschier, J.; Wood, W. I.; Goralka, T. M.; Wilson, K. L.; Chen, E. Y.; Eaton, D. H.; Vehar,
 Nature 312, 326-330, 1984
 A:Title: Characterization of the human factor VIII gene.
 A:Reference number: A56196; MUID:85061547; PMID:6438525
 A:Contents: annotation; introns
 R;McMullen, B. A.; Fujikawa, K.; Davie, E. W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o
 A:Reference number: A56216; MUID:95338127; PMID:7613471
 A:Contents: annotation; disulfide bonds
 R;Kjalk, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A:Reference number: S63527; MUID:96163459; PMID:8575434
 A:Accession: S63527
 A:Molecule type: protein
 A:Residues: 733-752;753-759 <K>
 R;Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A:Reference number: S66445; MUID:96048024; PMID:7556150
 A:Accession: S66445
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1668-1685 <LN>
 C:Comment: Factor VIII is activated by factor xa and thrombin, but prolonged exposure pr
 C:Genetics: GDB:F8C
 A:Gene: GDB:F8C
 A:Cross-references: GDB:119124; OMIM:306700
 A:Map position: Xq28-Xq28
 A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C:Function:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
 F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>
 F:20-356/Domain: A1 <DA1>
 F:23-348/Domain: ferroxidase repeat homology <FO1>
 F:392-759/Domain: A2 <DA2>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:760-1667/Domain: B <DB0>
 F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>
 F:1709-2038/Domain: A3 <DA3>
 F:1716-2038/Domain: ferroxidase repeat homology <FO3>
 F:2039-2191/Domain: C1 <DC1>
 F:2192-2351/Domain: discoidin I amino-terminal homology <DNI>
 F:2192-2345/Domain: C2 <DC2>
 F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #s;
 F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicts
 F:365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experime
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:414,426/Binding site: sulfate (Tyr) (covalent) #status experime
 F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experime
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:1740-1741/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:2193-2345/Disulfide bonds: #status predicted

Query Match 95.7%; Score 154; DB 1; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 6.9e-12;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPEPQLMKNEEAEDYDDDLT 30
 |||||
 DB 341 AYVKVDSCEPEPQLMKNEEAEDYDDDLT 370

RESULT 2
 A47004
 coagulation factor VIII precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 C:Accession: A47004
 R:Elder, B.; Iakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993

A:Title: Sequence of the murine factor VIII cDNA.
 A:Reference number: A47004; MUID:93300511; PMID:8314577
 A:Accession: A47004
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2319 <ELD>
 A:Cross-references: GB:L05573; NID:g192456; PIDN:AAA37385.1; PID:g192457
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:23-349/Domain: ferroxidase repeat homology <FO1>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:1686-2006/Domain: ferroxidase repeat homology <FO3>
 F:2007-2156/Domain: discoidin I amino-terminal homology <DNI>
 F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 70.5%; Score 113.5; DB 2; Length 2319;
 Best Local Similarity 80.0%; Pred. No. 1.5e-06;
 Matches 24; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 AYVKVDSCEPEPQLMK-NNEEAEDYDDDL 29
 |||||
 DB 342 AYVKVDSCEPEPQLMKNEEAEDYDDDL 371

RESULT 3
 T42763
 coagulation factor VIII precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42763
 R:Lollar, P.
 Submitted to the EMBL Data Library, August 1996
 A:Reference number: Z22269
 A:Accession: T42763
 A>Status: preliminary; translated from GB/EMBL/DDDBJ
 A:Molecule type: mRNA
 A:Residues: 1-2133 <LOL>
 A:Cross-references: EMBL:049517; NID:g1511633; PID:g1511634; PIDN:AAB06705.1
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A;
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
 F:23-349/Domain: ferroxidase repeat homology <FOX1>
 F:402-730/Domain: ferroxidase repeat homology <FOX2>
 F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

100

A:Molecule type: DNA
A:Residues: 1-537 <ROW>
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378319; PIDN:AAD03902.1
C:Genetics:
A:Genome: plasmid pNL1
A:Note: orf468
C:Superfamily: Sphingomonas aromaticivorans hypothetical protein 468

Query Match 32.3%; Score 52; DB 2; Length 537;
Best Local Similarity 40.0%; Pred. No. 44;
Matches 10; Conservative 6; Mismatches 5; Indels 4; Gaps 1;

QY 9 PEPOLIMKNEEAED----YDDDL 29
Db 499 PVEPEALVPDERVEDAALEYDDV 523

RESULT 9
I64847
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jun-2000
C:Accession: I64847
R:McDonald, F.J.; Snyder, P.M.; Price, M.P.; Welsh, M.J.
A:Title: Cloning and expression of the beta and gamma subunits of the human epithelial sodium channel gamma subunit - human
A:Reference number: I51915
A:Accession: I64847
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-649 <RES>
A:Cross-references: GB:L36592; NID:g845511; PIDN:AA75460.1; PID:g845512
C:Genetics:
A:Gene: gamma hENaC
C:Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I repeat

Query Match 32.3%; Score 52; DB 2; Length 649;
Best Local Similarity 37.9%; Pred. No. 54;
Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 AYKVQSCPEEPOLIMKNEEAEDYDDDL 29
Db 574 AWKQAPPCPEAPRSPQGDNPALDIDDDL 602

RESULT 10
I38204
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jun-2000
C:Accession: I38204; I39185
R:Voilley, N.; Bassilana, F.; Migon, C.; Merscher, S.; Mattei, M.G.; Carle, G.F.; Lazdun
A:Title: Cloning, chromosomal localization and physical linkage of the beta and gamma subunits of the human amiloride-sensitive sodium channel, nonvoltage-gated 1, gamma chain
A:Reference number: I38203; MUID:96039270; PMID:7490094
A:Accession: I38204
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-649 <RES>
A:Cross-references: EMBL:X87160; NID:g1004272; PIDN:CAA60633.1; PID:g1004273
R:Hansson, J.H.; Nelson-Williams, C.; Suzuki, H.; Schild, L.; Shimkets, R.; Lu, Y.; Cane
A:Title: Hypertension caused by a truncated epithelial sodium channel gamma subunit: genetic basis of a form of familial hypertension
A:Reference number: I39185; MUID:96021037; PMID:7550319
A:Accession: I39185
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 524-649 <RE2>
A:Cross-references: EMBL:U35630; NID:g1017735; PIDN:AAC50217.1; PID:g1017736
C:Genetics:
A:Gene: GDB:SCNN1G
A:Cross-references: GDB:568759; OMIM:600761

A:Map position: 16p12.2-16p12.1
C:Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I repeat

Query Match 32.3%; Score 52; DB 2; Length 649;
Best Local Similarity 37.9%; Pred. No. 54;
Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 AYKVQSCPEEPOLIMKNEEAEDYDDDL 29
Db 574 AWKQAPPCPEAPRSPQGDNPALDIDDDL 602

RESULT 11
A55969

FP21 protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 08-Dec-2000
C:Accession: A55969
R:Kozarov, E.; van der Wel, H.; Field, M.; Gritsalis, M.; Brown Jr., R.D.; West, C.M.
A:Title: Characterization of FP21, a cytosolic glycoprotein from Dictyostelium.
A:Reference number: A55969; MUID:95155385; PMID:7852383
A:Accession: A55969
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-162 <KQZ>
A:Cross-references: GB:U18063; NID:g639923; PIDN:AAA67888.1; PID:g639924
C:Superfamily: human S-phase kinase-associated protein 1A

Query Match 31.7%; Score 51; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 3 VKVDSCEPPEPOLIMKNEEAED 24
Db 137 IKNDFTPEEEQIRKENWCED 158

RESULT 12
C84476

Hypothetical protein At2g06220 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84476
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84476
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-428 <STO>
A:Cross-references: GB:AE002093; NID:g4567291; PIDN:AAD23703.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g06220
A:Map position: 2

Query Match 31.7%; Score 51; DB 2; Length 428;
Best Local Similarity 45.8%; Pred. No. 47;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 6 DSCPEEPOLIMKNEEAEDYDDDL 29
Db 128 DSGPEDEVTCKENEDEAGNDSL 151

RESULT 13
A97522

Ribonuclease E, RNase E (AF259270) [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:22 ; Search time 9.25 Seconds
(Without alignments)
152.519 Million cell updates/sec

Title: FVIII_ARG355I
Perfect score: 161
Sequence: 1 AYVKVDSCEPQLIMKNNRAEDYDDDLT 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	154	95.7	2351	1	F88_HUMAN
2	113.5	70.5	2319	1	F88_MOUSE
3	90.5	56.2	2133	1	F88_PIG
4	53	32.9	6632	1	UN89_CAEEL
5	52	32.3	649	1	SCAG_HUMAN
6	51.5	32.0	520	1	CET1_CANAL
7	51	31.7	162	1	FP21_DICDI
8	51	31.7	570	1	GRAU_DROME
9	50	31.1	475	1	Y409_HAEIN
10	50	31.1	1170	1	TSPI1_HUMAN
11	50	31.1	1170	1	TSPI1_MOUSE
12	49.5	30.7	350	1	ITP2_MOUSE
13	49.5	30.7	736	1	EF2_AERPE
14	49	30.4	400	1	GTR3_RABIT
15	48	29.8	675	1	VPS5_YEAST
16	48	29.8	682	1	NPH1_CAEEL
17	48	29.8	704	1	KPC1_CAEEL
18	48	29.8	734	1	EF2_DESMO
19	48	29.8	746	1	GYP7_YEAST
20	47.5	29.5	167	1	YJG5_YEAST
21	47.5	29.5	1906	1	DICE_MOUSE
22	47	29.2	104	1	GTR3_PIG
23	47	29.2	337	1	Y66L_SYNY3
24	47	29.2	495	1	GTR3_CANFA
25	47	29.2	496	1	GTR3_HUMAN
26	47	29.2	524	1	VGLG_RABVP
27	47	29.2	554	1	DHAB_SALTY
28	47	29.2	735	1	EF2_SULSO
29	47	29.2	886	1	YKRI_CAEEL
30	47	29.2	1170	1	TSPI1_BOVIN
31	47	29.2	1386	1	ZAP3_MOUSE
32	46.5	28.9	387	1	PAB_PEPMA
33	46	28.6	114	1	RLA1_HUMAN

34	46	28.6	338	1	TAP4_HUMAN	O01664 homo sapien
35	46	28.6	426	1	VG4_BPFID	P03664 bacterioph
36	46	28.6	494	1	GTR3_BOVIN	P58352 bos taurus
37	46	28.6	494	1	GTR3_SHEEP	P47843 ovis aries
38	46	28.6	494	1	SYK_SULSO	P95970 sulfolobus
39	46	28.6	522	1	IKAR_ONCMY	O13089 oncorhynch
40	46	28.6	736	1	EF2_SULTO	O975h5 sulfolobus
41	46	28.6	770	1	A4_PIG	P79307 s anyloid b
42	46	28.6	834	1	MSH5_HUMAN	O43196 homo sapien
43	46	28.6	918	1	IMB2_YEAST	P38217 saccharomyc
44	46	28.6	1173	1	TSPI_XENLA	P35448 xenopus lae
45	46	28.6	2175	1	HMCU_DROME	P10180 drosophila

ALIGNMENTS

RESULT 1

FA8_HUMAN	FA8_HUMAN	STANDARD;	PRT; 2351 AA.
AC P00451;			
DT 21-JUL-1986 (Rel. 01, Created)			
DT 21-JUL-1986 (Rel. 01, Last sequence update)			
DT 15-SEP-2003 (Rel. 42, Last annotation update)			
DE Coagulation factor VIII precursor (Procoagulant component)			
DE (Antihemophilic factor) (AHF).			
GN F8 OR F8C.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=86081164; PubMed=3935400;			
RA Triett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,			
RA Hartog K., Kuo C.H., Masiaz F.R., Merryweather J.P., Najarian R.,			
RA Pachi C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,			
RA Ordea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,			
RA Nordfang O., Ezban M.;			
RT "Characterization of the polypeptide composition of human factor			
VIII:C and the nucleotide sequence and expression of the human kidney			
cdna."			
RL DNA 4:333-349(1985).			
RN [2]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=85061548; PubMed=6438526;			
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,			
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,			
RA Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;			
RT "Expression of active human factor VIII from recombinant DNA clones."			
Nature 312:330-337(1984).			
RN [3]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=85061550; PubMed=6438528;			
RA Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,			
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,			
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,			
RA Hewick R.M.;			
RT "Molecular cloning of a cDNA encoding human antihaemophilic factor."			
Nature 312:342-347(1984).			
RN [4]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=93265012; PubMed=1303178;			
RA Gitschier J., Wood W.I.;			
RT "Sequence of the exon-containing regions of the human factor VIII			
gene."			
RL Hum. Mol. Genet. 1:199-200(1992).			
RN [5]			
RP SEQUENCE OF 2064-2070 FROM N.A.			
RA de Water N.S., Williams R., Browett P.J.;			
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RN [6]			
RP SULFATION OF TYR-1699.			

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor.";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP VARIANT HEMA CYS-391.
 RA MEDLINE=92207952; PubMed=1534716;
 RX Pittman D.D., Wang J.H., Kaufman R.J.;
 RA "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RA MEDLINE=95200924; PubMed=7893714;
 RX Gilbert G.E., Baleja J.D.;
 RA "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=95243332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT HEMA GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac.";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT HEMA PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT HEMA GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT HEMA GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Yousoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RL A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fuicher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT HEMA LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT HEMA LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HEMA HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Aral M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Aral M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANTS HEMA GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS HEMA PHE-1699 AND CYS-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:65-71(1990).
 RN [25]
 RP VARIANTS HEMA CYS-1728 AND ASP-1941.
 RX MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

FT	DOMAIN	1855	2008	PLASTOCYANIN- LIKE 6.
FT	DOMAIN	2061	2156	F5/8 TYPE C 1.
FT	DOMAIN	2161	2313	F5/8 TYPE C 2.
FT	SITE	351	392	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	759	760	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	1678	1679	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	1324	1325	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	SITE	1640	1641	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	MOD_RES	367	367	SULFATION (BY SIMILARITY).
FT	MOD_RES	737	737	SULFATION (BY SIMILARITY).
FT	MOD_RES	738	738	SULFATION (BY SIMILARITY).
FT	MOD_RES	742	742	SULFATION (BY SIMILARITY).
FT	MOD_RES	1669	1669	SULFATION (REQUIRED FOR VWF BINDING) (BY SIMILARITY).
FT	MOD_RES	1687	1687	SULFATION (BY SIMILARITY).
FT	DISULFID	173	199	PROBABLE.
FT	DISULFID	547	573	PROBABLE.
FT	DISULFID	1819	1845	PROBABLE.
FT	DISULFID	2008	2156	BY SIMILARITY.
FT	DISULFID	2161	2313	BY SIMILARITY.
FT	CARBOHYD	61	61	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	423	423	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	601	601	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	958	958	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1015	1015	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1022	1022	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1026	1026	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1044	1044	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1076	1076	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1087	1087	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1136	1136	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1161	1161	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1192	1192	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1255	1255	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1268	1268	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1273	1273	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1274	1274	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1302	1302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1316	1316	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1340	1340	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1378	1378	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1797	1797	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	2105	2105	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	2319 AA;	266148 MW;	FO054DE051DB2A01 CRC64;
Query Match			70.5%;	Score 113.5; DB 1; Length 2319;
Best Local Similarity			80.0%;	Pred. No. 4e-07;
Matches 24;	Conservative		0;	Mismatches 5; Indels 1; Gaps 1;
QY	1	AYVKVDSCEPPQOLIMK-NNEAEYDDDL 29		
Db	342	AYVKVDSCEPESQWQKNNNEEDYDDDL 371		
RESULT 3				
FA8_PIG				
ID	FA8_PIG	STANDARD;	PRT;	2133 AA.
PI2263;				
01-OCT-1989				
01-NOV-1997				
15-SEP-2003				
Coagulation factor VIII precursor (Procoagulant component).				
F8 OR CF8.				
OS	Sus scrofa (pig).			
Eukaryota; Metazoa;				
Mammalia; Eutheria;				
Cetartiodactyla; Suina;				
Suidae; Sus.				
NCBI_TaxID=9823;				
[1]				
SEQUENCE FROM N.A.				

Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

[2]
 RP SEQUENCE OF 705-1573 FROM N.A.
 RX MEDLINE=86287369; PubMed=3016730;
 RA Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
 RA Kaufman R.J.;
 RT "A large region (approximately equal to 95 kDa) of human factor VIII
 is dispensable for in vitro procoagulant activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
 RN [3]

RP SEQUENCE OF 392-759 FROM N.A.
 RX MEDLINE=94179260; PubMed=7510693;
 RA Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
 RT "Elimination of a major inhibitor epitope in factor VIII.";
 RL J. Biol. Chem. 269:8639-8641(1994).
 RN [3]

CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 ACTIVATED FORM, FACTOR XA.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U49517; AAB06705.1; -
 DR PIR: A25945; A25945.
 DR PIR: T42763; T42763.
 DR HSP: P00451; ICRF.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS00022; FA58C_3; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 KW Signal; Glycoprotein; Sulfation.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 2133 COAGULATION FACTOR VIII.
 FT DOMAIN 20 357 F5/8 TYPE A 1.
 FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 207 357 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 399 730 F5/8 TYPE A 2.
 FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 760 1599 B.
 FT DOMAIN 1495 1822 F5/8 TYPE A 3.
 FT DOMAIN 1495 1659 PLASTOCYANIN-LIKE 5.
 FT DOMAIN 1669 1822 PLASTOCYANIN-LIKE 6.
 FT DOMAIN 1822 1970 F5/8 TYPE C 1.
 FT DOMAIN 1975 2127 F5/8 TYPE C 2.
 FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1449 1450 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1490 1491 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 737 737 SULFATION (BY SIMILARITY).
 FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
 FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
 FT DISULFID 173 199 PROBABLE.
 FT DISULFID 547 573 PROBABLE.
 FT DISULFID 1633 1659 PROBABLE.
 FT DISULFID 1822 1970 BY SIMILARITY.
 FT DISULFID 1975 2127 BY SIMILARITY.
 FT CARBOHYD 233 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 929 929 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 985 985 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1025 1025 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1111 1111 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1181 1181 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1208 1208 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1245 1245 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1265 1265 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1335 1335 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1408 1408 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1611 1611 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1919 1919 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 713 713 N -> M (IN REF. 2).
 FT CONFLICT 734 734 I -> T (IN REF. 2).
 FT CONFLICT 792 792 E -> F (IN REF. 2).
 FT CONFLICT 1133 1133 E -> F (IN REF. 2).
 FT CONFLICT 1191 1191 I -> L (IN REF. 2).
 FT CONFLICT 1209 1209 R -> F (IN REF. 2).
 FT CONFLICT 1437 1437 C -> G (IN REF. 2).
 FT CONFLICT 1456 1456 F -> R (IN REF. 2).
 FT CONFLICT 1539 1539 F -> R (IN REF. 2).
 FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
 SQ SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;
 Query Match 56.2%; Score 90.5; DB 1; Length 2133;
 Best Local Similarity 65.5%; Pred. No. 0.00047;
 Matches 19; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
 QY 1 AYVVDSCPEEPQQLMKNEEAEDYDDL 29
 DB 342 AHRVSCAEEPPQLRRKADEE-EDYDDL 369
 RESULT 4
 UN89_CAEEL STANDARD; PRT: 6632 AA.
 ID UN89_CAEEL
 AC 001761; Q17362;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
 GN UNC-89 OR C09D1.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=Bristol N2;
 RX MEDLINE=96180278; PubMed=8603916;
 RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
 RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
 assembly, encodes a giant modular protein composed of Ig and signal
 transduction domains.";
 RL J. Cell Biol. 132:835-848(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Du Z., Le T.I., Willson R.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Structural component of the muscle M-line. Myoflament
 lattice assembly begins with positional cues laid down in the
 basement membrane and muscle cell membrane. UNC-89 responds to
 these signals, localizes, and then participates in assembling an
 M-line.
 CC -!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
 CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.

CC CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
 CC CC -1- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
 CC CC -1- SIMILARITY: Contains 1 PH domain.
 CC CC -1- SIMILARITY: Contains 5 RCD domain.
 CC CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL; U33058; AAB00542.1; -
 CC CC EMBL; AF003131; AAB54132.2; -
 CC CC PDB; 1PHQ; 20-DEC-00
 CC CC WormPep; C09D1.1; CE30426.
 CC CC InterPro; IPR003961; FN-III.
 CC CC InterPro; IPR007110; Ig-like.
 CC CC InterPro; IPR003598; Ig_C2.
 CC CC InterPro; IPR003006; Ig_MHC.
 CC CC InterPro; IPR001849; PH.
 CC CC InterPro; IPR007850; RCD.
 CC CC InterPro; IPR000219; RhoGEF.
 CC CC InterPro; IPR001452; SH3.
 CC CC Pfam; PF00041; f03; 1.
 CC CC Pfam; PF00047; ig; 47.
 CC CC Pfam; PF00169; PH; 1.
 CC CC Pfam; PF05177; RCD; 5.
 CC CC Pfam; PF00621; RhoGEF; 1.
 CC CC Pfam; PF00018; SH3; 1.
 CC CC SMART; SM00408; IgC2; 23.
 CC CC SMART; SM00325; RhoGEF; 1.
 CC CC SMART; SM00326; SH3; 1.
 CC CC PROSITE; PS50010; DH2; 1.
 CC CC PROSITE; PS50835; IG-LIKE; 49.
 CC CC PROSITE; PS50003; PH_DOMAIN; 1.
 CC CC PROSITE; PS50002; SH3; 1.
 CC CC Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
 CC CC 3D-structure.
 CC CC DOMAIN 63 127
 CC CC DOMAIN 152 330
 CC CC PH.
 CC CC IG-LIKE C2-TYPE 1.
 CC CC IG-LIKE C2-TYPE 2.
 CC CC IG-LIKE C2-TYPE 3.
 CC CC IG-LIKE C2-TYPE 4.
 CC CC IG-LIKE C2-TYPE 5.
 CC CC IG-LIKE C2-TYPE 6.
 CC CC THR-RICH.
 CC CC RCD 1.
 CC CC RCD 2.
 CC CC RCD 3.
 CC CC RCD 4.
 CC CC RCD 5.
 CC CC RCD 5.
 CC CC IG-LIKE C2-TYPE 7.
 CC CC IG-LIKE C2-TYPE 8.
 CC CC IG-LIKE C2-TYPE 9.
 CC CC IG-LIKE C2-TYPE 10.
 CC CC IG-LIKE C2-TYPE 11.
 CC CC IG-LIKE C2-TYPE 12.
 CC CC IG-LIKE C2-TYPE 13.
 CC CC IG-LIKE C2-TYPE 14.
 CC CC IG-LIKE C2-TYPE 15.
 CC CC IG-LIKE C2-TYPE 16.
 CC CC IG-LIKE C2-TYPE 17.
 CC CC IG-LIKE C2-TYPE 18.
 CC CC IG-LIKE C2-TYPE 19.
 CC CC IG-LIKE C2-TYPE 20.
 CC CC IG-LIKE C2-TYPE 21.
 CC CC IG-LIKE C2-TYPE 22.
 CC CC IG-LIKE C2-TYPE 23.
 CC CC 3567

DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.
 DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
 DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
 DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.
 DOMAIN 4109 4201 IG-LIKE C2-TYPE 28.
 DOMAIN 4212 4297 IG-LIKE C2-TYPE 29.
 DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.
 DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.
 DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.
 DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.
 DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.
 DOMAIN 4873 4961 IG-LIKE C2-TYPE 35.
 DOMAIN 4965 5057 IG-LIKE C2-TYPE 36.
 DOMAIN 5067 5160 IG-LIKE C2-TYPE 37.
 DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.
 DOMAIN 5277 5366 IG-LIKE C2-TYPE 39.
 DOMAIN 5383 5472 IG-LIKE C2-TYPE 40.
 DOMAIN 5487 5578 IG-LIKE C2-TYPE 41.
 DOMAIN 5595 5685 IG-LIKE C2-TYPE 42.
 DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.
 DOMAIN 5815 5904 IG-LIKE C2-TYPE 44.
 DOMAIN 5925 6014 IG-LIKE C2-TYPE 45.
 DOMAIN 6038 6130 IG-LIKE C2-TYPE 46.
 DOMAIN 6150 6239 IG-LIKE C2-TYPE 47.
 DOMAIN 6275 6368 FIBRONECTIN TYPE-III.
 DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.
 DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.
 DISULFID 568 621 POTENTIAL.
 DISULFID 2908 2975 POTENTIAL.
 DISULFID 3015 3065 POTENTIAL.
 DISULFID 3707 3759 POTENTIAL.
 DISULFID 3826 3890 POTENTIAL.
 DISULFID 5092 5157 POTENTIAL.
 DISULFID 5298 5350 POTENTIAL.
 DISULFID 5508 5560 POTENTIAL.
 DISULFID 5616 5669 POTENTIAL.
 DISULFID 5722 5784 POTENTIAL.
 DISULFID 5836 5901 POTENTIAL.
 DISULFID 5946 5998 POTENTIAL.
 DISULFID 6036 6171 POTENTIAL.
 DISULFID 6421 6486 POTENTIAL.
 DISULFID 2137 2137 A -> P (IN REF. 1).
 CONFLICT 2137 2137 AKA -> PKP (IN REF. 1).
 CONFLICT 2245 2247 A -> P (IN REF. 1).
 CONFLICT 2258 2258 E -> G (IN REF. 1).
 CONFLICT 2284 2284 M -> I (IN REF. 1).
 CONFLICT 2297 2297 A -> G (IN REF. 1).
 CONFLICT 3531 3531 DAGEY -> RRRRI (IN REF. 1).
 CONFLICT 3884 3888 A -> V (IN REF. 1).
 CONFLICT 3929 3929 A -> P (IN REF. 1).
 CONFLICT 5134 5134 T -> S (IN REF. 1).
 CONFLICT 5145 5145 G -> A (IN REF. 1).
 CONFLICT 5185 5185 K -> N (IN REF. 1).
 CONFLICT 5199 5199 L -> F (IN REF. 1).
 CONFLICT 5202 5202 F -> L (IN REF. 1).
 CONFLICT 5213 5213 A -> G (IN REF. 1).
 CONFLICT 6178 6178 K -> E (IN REF. 1).
 CONFLICT 6268 6268
 SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960B89 CRC64;
 Query Match 32.9%; Score 53; DB 1; Length 6632;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 2 YKVDSCPEEPQLIMKNEE 21
 I:: I:: I:: I:: I::
 Db 178 YIKVLDOPEPEAVKKNREE 197
 RESULT 5
 SCAG_HUMAN
 ID SCAG_HUMAN STANDARD; PRT: 649 AA.
 AC P51170; P78437; Q93023; Q93024; Q93025; Q93026; Q93027;
 DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Amiloride-sensitive sodium channel gamma-subunit (Epithelial Na+
 DE channel gamma subunit) (Gamma ENaC) (Nonvoltage-gated sodium channel
 DE 1 gamma subunit) (SCN3G) (Gamma NACH).
 GN SCN3G.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=96039270; PubMed=7490094;
 RA Volleijn N., Bassilana F., Mignon C., Merscher S., Mattei M.-G.,
 RA Carle G.F., Lazdunski M., Barbry P.;
 RT "Cloning, chromosomal localization, and physical linkage of the beta
 RT and gamma subunits (SCN3B and SCN3G) of the human epithelial
 RT amiloride-sensitive sodium channel.";
 RL Genomics 28:560-565(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95282837; PubMed=7762608;
 RA McDonald F.J., Snyder P.M., Price M.P., Welsh M.J.;
 RA "Cloning and expression of the beta- and gamma-subunits of the human
 RT epithelial sodium channel.";
 RL Am. J. Physiol. 268:C1157-C1163(1995).
 [3]
 RP SEQUENCE OF 1-43; 332-392; 433-463 AND 499-515 FROM N.A.
 RX MEDLINE=96421599; PubMed=8924247;
 RA Thomas C.P., Doggett N.A., Fisher R., Stokes J.B.;
 RA "Genomic organization and the 5' flanking region of the gamma subunit
 RT of the human amiloride-sensitive epithelial sodium channel.";
 RL J. Biol. Chem. 271:26062-26066(1996).
 CC -!- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE-SENSITIVE ION CHANNEL
 CC INHIBITED BY THE DIURETIC AMILORIDE. MEDIATES THE ELECTRODIFFUSION
 CC OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY)
 CC THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE
 CC REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS.
 CC ALSO PLAYS A ROLE IN TASTE PERCEPTION.
 CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA, ONE BETA AND ONE GAMMA
 CC SUBUNIT. A DELTA SUBUNIT CAN REPLACE THE ALPHA SUBUNIT.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- DISEASE: DEFECTS IN SCN3G ARE ONE OF THE CAUSES OF LIDDLE
 CC SYNDROME (PSEUDALDOSTERONISM), AN AUTOSOMAL DOMINANT FORM OF
 CC HYPERTENSION ASSOCIATED WITH HYPOKALEMIC ALKALOSIS. THE DISEASE IS
 CC CAUSED BY CONSTITUTIVE ACTIVATION OF THE RENAL EPITHELIAL SODIUM
 CC CHANNEL.
 CC -!- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNEL
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X87160; CAA60633.1; -
 DR EMBL; L36592; AAA75460.1; -
 DR EMBL; U48936; AAC50737.1; -
 DR EMBL; U53836; AAC50744.1; -
 DR EMBL; U53845; AAC50753.1; -
 DR EMBL; U53846; AAC50754.1; -
 DR EMBL; U53847; AAC50755.1; -
 DR EMBL; U53850; AAC50758.1; -
 DR EMBL; U53852; AAC50760.1; -
 DR PIR; I38204; I38204.
 DR PIR; I64847; I64847.
 DR Genew; HGNC:10602; SCN3G.
 DR MIM; 600761; -

DR MIM; 177200; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0015280; P:amiloride-sensitive sodium channel activity; TAS.
 DR GO; GO:0007588; P:excretion; TAS.
 DR GO; GO:0006814; P:sodium ion transport; TAS.
 DR InterPro; IPR004724; EnaC.
 DR InterPro; IPR001873; Na+channel_ASC.
 DR Pfam; PF00858; ASC; 1.
 DR PRINTS; PRO1078; AMINACHANNEL.
 DR TIGRFAMs; TIGR00859; ENaC; 1.
 DR PROSITE; PS01206; ASC; 1.
 KW Ionic channel; Transmembrane; Ion transport; Glycoprotein;
 KW Polymorphism.
 FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 56 76 POTENTIAL.
 FT DOMAIN 77 541 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 542 562 POTENTIAL.
 FT DOMAIN 563 649 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 49 49 G -> C (IN dbSNP:5733).
 FT VARIANT 183 183 G -> S (IN dbSNP:5736).
 FT CONFLICT 178 178 R -> W (IN REF. 2).
 FT CONFLICT 339 339 F -> S (IN REF. 1).
 FT CONFLICT 350 350 A -> T (IN REF. 1).
 FT CONFLICT 369 369 Y -> S (IN REF. 1).
 FT CONFLICT 375 375 D -> G (IN REF. 1).
 FT CONFLICT 458 458 R -> S (IN REF. 2 AND 3).
 FT CONFLICT 461 463 EWT -> DGH (IN REF. 3).
 FT CONFLICT 502 502 P -> A (IN REF. 1).
 FT CONFLICT 614 614 A -> S (IN REF. 2).
 SQ SEQUENCE 649 AA; 74365 MW; 02DE31178A74D2FF CRC64;
 Query Match 32.3%; Score 52; DB 1; Length 649;
 Best Local Similarity 37.9%; Pred. No. 21;
 Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
 QY 1 AYVKVDSCEEPQQLMKNEEAEDYDDL 29
 I: : | | | : | | | | |
 Db 574 AWKAPPCPEAPSPQGDNPALDIDDL 602
 RESULT 6
 CETL_CANAL
 ID CETL_CANAL STANDARD; PRT; 520 AA.
 AC 093803;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE mRNA capping enzyme beta subunit (Polynucleotide 5'-triphosphatase)
 DE (EC 3.1.3.33) (mRNA 5'-triphosphatase) (IPase).
 GN CETL.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IFO 1060;
 RX MEDLINE=98427288; PubMed=9755857;
 RA Yamada-Okabe T., Mio T., Matsui M., Kashima Y., Arisawa M.,
 RA Yamada-Okabe H.;
 RT "Isolation and characterization of the Candida albicans gene for mRNA
 RT 5'-triphosphatase: association of mRNA 5'-triphosphatase and mRNA 5'-
 RT guanylyltransferase activities is essential for the function of mRNA
 RT 5'-capping enzyme in vivo.";
 RL FEBS Lett. 435:49-54(1998).
 CC -!- FUNCTION: FIRST STEP OF M-RNA CAPPING. CONVERTS THE 5'-

CC TRIPHOSPHATE END OF A NASCENT MENA CHAIN INTO A DIPHOSPHATE END.
 CC -1- CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
 CC polynucleotide + phosphate.
 CC -1- COFACTOR: REQUIRES DIVALENT IONS (BY SIMILARITY).
 CC -1- SUBUNIT: THE M-RNA CAPPING ENZYME IS COMPOSED OF TWO SEPARATE
 CC CHAINS ALPHA AND BETA, RESPECTIVELY A MRNA GUANYLYLTRANSFERASE AND
 CC AN RNA 5'-TRIPHOSPHATASE.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FUNGAL TPASE FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB016242; BAA33965.1; -
 CC InterPro; IPR004206; mRNA_tripase.
 CC Pfam; PF02940; mRNA_tripase; 1
 CC Hydrolase; mRNA processing; mRNA capping; Nuclear protein.
 CC DOMAIN 92 95 POLY-SER.
 CC SEQUENCE 520 AA; 58791 MW; 5C0690F347E13EEC CRC64;
 CC
 CC Query Match 32.0%; Score 51.5; DB 1; Length 520;
 CC Best Local Similarity 45.2%; Pred. No. 19;
 CC Matches 14; Conservative 2; Mismatches 4; Indels 11; Gaps 2;
 CC
 CC QY 5 VDSCEP-----EPQLIMKNEEAFDYDDLT 30
 CC ||||| ||| : |||||
 CC 164 VDSAPFPKPKKEQPVF-----DQDDDLT 188
 CC
 CC RESULT 7
 CC FP21_DICDI STANDARD; PRT; 162 AA.
 CC ID FP21_DICDI
 CC AC P52285; Q94505;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Glycoprotein FP21 precursor.
 CC GN (FPA1A OR FP21A OR FPA1 OR SKP1A) AND (FPA1B OR FP21B OR FPA2 OR
 CC SKP1B).
 CC OS Dictyostelium discoideum (slime mold).
 CC OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostellium.
 CC OX NCBI_TaxID=44689;
 CC RN [1]
 CC RP SEQUENCE FROM N.A. (FP21A), AND SEQUENCE OF 31-43.
 CC RC STRAIN=AX3;
 CC RX MEDLINE=95155385; PubMed=7852383;
 CC RA Kozartov E., van der Wel H., Field M., Gritzali M., Brown R.D. Jr.,
 CC West C.M.;
 CC RT "Characterization of FP21, a cytosolic glycoprotein from
 CC Dictyostellium.";
 CC RL J. Biol. Chem. 270:3022-3030(1995).
 CC [2]
 CC RP SEQUENCE FROM N.A. (FP21A AND FP21B).
 CC RC STRAIN=AX3;
 CC RX MEDLINE=98038971; PubMed=9373134;
 CC RA West C.M., Kozarov E., Teng-Ummay P.;
 CC RT "The cytosolic glycoprotein FP21 of Dictyostelium discoideum is
 CC encoded by two genes resulting in a polymorphism at a single amino
 CC acid position.";
 CC RL Gene 200:1-10(1997).
 CC [3]
 CC RP SEQUENCE FROM N.A. (FP21B).
 CC RC STRAIN=AX4;
 CC RX Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
 CC Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 CC Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A.,
 CC Noegel A.A.;
 CC RT "Sequence and analysis of chromosome 2 of Dictyostellium.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 4-38.
 RX MEDLINE=92250608; PubMed=1577798;
 RA Gonzalez-Vanes B., Cicero J.M., Brown R.D. Jr., West C.M.;
 RT "Characterization of a cytosolic fucosylation pathway in
 CC Dictyostellium.";
 CC RL J. Biol. Chem. 267:9595-9605(1992).
 RN [5]
 RP CARBOHYDRATE-LINKAGE SITE PRO-143, AND MASS SPECTROMETRY.
 RX MEDLINE=98325031; PubMed=9660787;
 RA Teng-Ummay P., Morris H.R., Dell A., Panico M., Paxton T., West C.M.;
 RT "The cytoplasmic F-box binding protein SKP1 contains a novel
 CC pentasaccharide linked to hydroxyproline in Dictyostellium.";
 CC J. Biol. Chem. 273:18242-18249(1998).
 CC -1- SUBUNIT: Multi-protein complex (SCF) with cullin and F-box-
 CC containing protein. Capable of undergoing aggregation.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- PTM: O-linked glycan consists of linear Gal-Gal-Fuc-Gal-GlcNAc.
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES CODING FOR FP21, THEY ONLY
 CC DIFFER IN A SINGLE POSITION IN THE CODING REGION.
 CC -1- SIMILARITY: TO YEAST CBF3D AND MAMMALIAN SKP1.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U18063; AAA67888.1; -
 CC EMBL; U73685; AAB88389.1; -
 CC EMBL; U73686; AAB88390.1; -
 CC EMBL; AC116030; AAL92996.1; -
 CC PIR; A55969; A55969.
 CC GlycoSuiteDB; P52285; -
 CC DictyDb; DD02045; fpa1A.
 CC DictyDb; DD07???; fpa1B.
 CC InterPro; IPR001232; Skp1.
 CC Pfam; PF01466; Skp1; 1.
 CC Pfam; PF03931; Skp1_P02; 1.
 CC SMART; SM00512; Skp1; 1.
 CC Glycoprotein; Hydroxylation.
 CC PROPEP 1
 CC CHAIN 4 162 GLYCOPROTEIN FP21.
 CC MOD_RES 143 143 HYDROXYLATION (PARTIAL).
 CC CARBOHYD 143 143 O-LINKED (GLCNAC).
 CC /FTID=CAR_000228.
 CC VARIANT 39 39 S->A (IN FP21B).
 CC SEQUENCE 162 AA; 18718 MW; 622D38D90CF9DBED CRC64;
 CC
 CC Query Match 31.7%; Score 51; DB 1; Length 162;
 CC Best Local Similarity 50.0%; Pred. No. 6.1;
 CC Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 CC
 CC QY 3 VKVDSCEPQLIMKNEEAFD 24
 CC :| | | | :| | | | |
 CC Db 137 IKNDFTPEEEQIRKENEWCED 158
 CC
 CC RESULT 8
 CC GRAU_DROME STANDARD; PRT; 570 AA.
 CC ID GRAU_DROME
 CC AC Q90405; Q9W2N7;
 CC DT 15-SEP-2003 (Rel. 42, Created)
 CC DT 15-SEP-2003 (Rel. 42, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Transcription factor grauzone.
 CC GN GRAU OR CG3282.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical metalloprotease HI0409 (EC 3.4.24.-).
 GN HI0409.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -!- CORACTOR: Zinc (potential).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M37.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32724; AAC22068.1; -;
 DR PIR; F64151; F64151.
 DR MEROPS; M37.0PW; -;
 DR TIGR; HI0409; -;
 DR InterPro; IPR002886; Peptidase M37.
 DR Pfam; PF01551; Peptidase_M37; 1.
 KW Hypothetical protein; hydrolase; Metalloprotease; Zinc;
 KW Complete proteome. 348 ZINC (POTENTIAL).
 FT METAL 348
 SQ SEQUENCE 475 AA; 53255 NW; 143C10F92233939D CRC64;
 Query Match 31.1%; Score 50; DB 1; Length 475;
 Best Local Similarity 32.1%; Pred. No. 27;
 Matches 9; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 QY 2 YKVDSCPEPQLIMKNEEADYDDL 29
 Db 57 YHELTSPNENSTALQPDENATSYDEL 84
 RESULT 10
 TSPL_HUMAN
 ID TSPL_HUMAN STANDARD; PRT; 1170 AA.
 AC P07996; Q15667;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Thrombospondin 1 precursor.
 GN THBS1 OR TSPL OR TSP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Endothelial cells;
 RX MEDLINE=87057617; PubMed=2430973;
 RA Lawler J., Hynes R.O.;

RT "The structure of human thrombospondin, an adhesive glycoprotein with
 RT multiple calcium-binding sites and homologues with several different
 RT proteins.";
 RL J. Cell Biol. 103:1635-1648(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89139590; PubMed=2918029;
 RX Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;
 RT "Complete thrombospondin mRNA sequence includes potential regulatory
 RT sites in the 3' untranslated region.";
 RL J. Cell Biol. 108:729-736(1989).
 RN [3]
 RP SEQUENCE OF 1-397 FROM N.A.
 RC MEDLINE=87157592; PubMed=3030396;
 RX Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
 RA "Partial amino acid sequence of human thrombospondin as determined by
 RT analysis of cDNA clones: homology to malarial circumsporozoite
 RT proteins.";
 RL Biochemistry 25:8418-8425(1986).
 RN [4]
 RP SEQUENCE OF 1-374 FROM N.A.
 RC MEDLINE=86287276; PubMed=3461443;
 RX Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
 RA "Characterization of a cDNA encoding the heparin and collagen binding
 RT domains of human thrombospondin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
 RN [5]
 RP SEQUENCE OF 1-166 FROM N.A.
 RC MEDLINE=89291870; PubMed=2544587;
 RX Laherty C.D., Gierman T.M., Dixit V.M.;
 RA "Characterization of the promoter region of the human thrombospondin
 RT gene. DNA sequences within the first intron increase transcription.";
 RL J. Biol. Chem. 264:11222-11227(1989).
 RN [6]
 RP SEQUENCE OF 1028-1170 FROM N.A.
 RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
 RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
 RC THR-450; TRP-498 AND THR-507.
 RX TISSUE=Platelet;
 RA MEDLINE=21125860; PubMed=11067851;
 RA Hofsteenge J., Howler K.G., Macek B., Hess D., Lawler J.,
 RA Mosher D.F., Peter-Katalinic J.;
 RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
 RT module.";
 RL J. Biol. Chem. 276:6485-6498(2001).
 RN [8]
 RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
 RX MEDLINE=22338361; PubMed=12450399;
 RA Howler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
 RT "Biophysical characterization, including disulfide bond assignments,
 RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
 RL Biochemistry 41:14329-14339(2002).
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
 CC V/beta-3 and alpha-IIb/beta-3.
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
 CC -!- SIMILARITY: Contains 1 WFCC domain.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

```

CC  EMBL; M25631; AAA36741.1; -
DR  EMBL; X04665; CAA28370.1; -
DR  EMBL; X14787; CAA32889.1; -
DR  EMBL; M14326; AAA61237.1; ALT_SEQ.
DR  EMBL; J04835; AAA61178.1; -
DR  EMBL; M99425; AAB59366.1; -
DR  PIR; A26155; TSHUP1.
DR  PDB; 1LSL; 18-DEC-02.
DR  GlycoSuiteDB; P07996; -.
DR  Genew; HGNC:11785; THBS1.
DR  MIM; 188060; -.
DR  GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR  GO; GO:0004871; F:signal transducer activity; TAS.
DR  GO; GO:0007275; P:development; TAS.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWFC_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF00093; vwc; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF02412; tsp_3; 8.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS00092; TSP1; 3.
DR  PROSITE; PS01208; VWFC; 1; 1.
DR  PROSITE; PS0184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal; 3D-structure.
FT  SIGNAL 1 18
FT  CHAIN 19 1170 THROMBOSPONDIN 1.
FT  DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
FT  DOMAIN 24 221 TSP N-TERMINAL.
FT  DOMAIN 316 373 VWFC.
FT  DOMAIN 379 429 TSP TYPE-1 1.
FT  DOMAIN 435 490 TSP TYPE-1 2.
FT  DOMAIN 492 547 TSP TYPE-1 3.
FT  DOMAIN 549 587 EGF-LIKE 1.
FT  DOMAIN 588 645 EGF-LIKE 2. CALCIUM-BINDING (POTENTIAL).
FT  DOMAIN 646 690 EGF-LIKE 3.
FT  DOMAIN 723 758 TSP TYPE-3 1.
FT  DOMAIN 759 781 TSP TYPE-3 2.
FT  DOMAIN 782 817 TSP TYPE-3 3.
FT  DOMAIN 818 840 TSP TYPE-3 4.
FT  DOMAIN 841 878 TSP TYPE-3 5.
FT  DOMAIN 879 914 TSP TYPE-3 6.
FT  DOMAIN 915 950 TSP TYPE-3 7.
FT  DOMAIN 951 1170 C-TERMINAL.
FT  SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
FT  SITE 926 970 INTERCHAIN (PROBABLE).
FT  DISULFID 274 274 INTERCHAIN (PROBABLE).
FT  DISULFID 391 423
FT  DISULFID 395 428
FT  DISULFID 447 484
FT  DISULFID 451 489
FT  DISULFID 462 474
FT  DISULFID 504 541
FT  DISULFID 508 546
FT  DISULFID 519 531
FT  DISULFID 551 562
FT  DISULFID 556 572
FT  DISULFID 575 586
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.

```

```

FT  DISULFID 592 608 BY SIMILARITY.
FT  DISULFID 599 617 BY SIMILARITY.
FT  DISULFID 620 644 BY SIMILARITY.
FT  DISULFID 650 663 BY SIMILARITY.
FT  DISULFID 657 676 BY SIMILARITY.
FT  DISULFID 678 689 BY SIMILARITY.
FT  DISULFID 705 713 BY SIMILARITY.
FT  DISULFID 718 738 BY SIMILARITY.
FT  DISULFID 754 774 BY SIMILARITY.
FT  DISULFID 777 797 BY SIMILARITY.
FT  DISULFID 813 833 BY SIMILARITY.
FT  DISULFID 836 856 BY SIMILARITY.
FT  DISULFID 874 894 BY SIMILARITY.
FT  DISULFID 910 930 BY SIMILARITY.
FT  DISULFID 946 1167 BY SIMILARITY.
FT  CARBOHYD 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 385 C-LINKED (MAN).
FT  CARBOHYD 394 O-LINKED (FUC. . .).
FT  CARBOHYD 438 /FTID=CAR_000206.
FT  CARBOHYD 441 /FTID=CAR_000207.
FT  CARBOHYD 450 /FTID=CAR_000208.
FT  CARBOHYD 498 /FTID=CAR_000209.
FT  CARBOHYD 507 /FTID=CAR_000210.
FT  CARBOHYD 507 /FTID=CAR_000211.
FT  CARBOHYD 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CONFLICT 84 T -> A (IN REF. 2; 3 AND 4).
FT  CONFLICT 523 T -> A (IN REF. 2).
SQ  SEQUENCE 1170 AA; 129412 MW; 69B3DE5AE3A395E CRC64;

Query Match 31.1%; Score 50; DB 1; Length 1170;
Best Local Similarity 40.7%; Pred. No. 74;
Matches 11; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

Qy 2 YKVDSCPEEPQLIMKNEEAEDYDD 28
Db 712 HCKKDCPNLF-----NSGQEDYDKD 732

RESULT 11
TSPL_MOUSE
ID TSPL_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.
RT "Characterization of mouse thrombospondin 2 sequence and expression
RT during cell growth and development.";

```

J. Biol. Chem. 267:3274-3281(1992).

[3]

SEQUENCE OF 1-490 FROM N.A.

MEDLINE=90375546; PubMed=2398070;

Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;

RT "Characterization of the mouse thrombospondin gene and evaluation of

RT the role of the first intron in human gene expression.";

J. Biol. Chem. 265:16691-16698(1990).

CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and

CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-

CC V/beta-3 and alpha-11b/beta-3.

CC -!- SUBUNIT: Homotrimer; disulfide-linked.

CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.

CC -!- SIMILARITY: Contains 1 WFEC domain.

CC -!- SIMILARITY: Contains 3 EGF-like domains.

CC -!- SIMILARITY: Contains 7 TSP type-1 domains.

CC -!- SIMILARITY: Contains 7 TSP type-3 domains.

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M62470; AAA50611.1; JOINED.

DR EMBL; M62450; AAA50611.1; JOINED.

DR EMBL; M62451; AAA50611.1; JOINED.

DR EMBL; M62452; AAA50611.1; JOINED.

DR EMBL; M62453; AAA50611.1; JOINED.

DR EMBL; M62454; AAA50611.1; JOINED.

DR EMBL; M62455; AAA50611.1; JOINED.

DR EMBL; M62456; AAA50611.1; JOINED.

DR EMBL; M62457; AAA50611.1; JOINED.

DR EMBL; M62458; AAA50611.1; JOINED.

DR EMBL; M62459; AAA50611.1; JOINED.

DR EMBL; M62460; AAA50611.1; JOINED.

DR EMBL; M62461; AAA50611.1; JOINED.

DR EMBL; M62462; AAA50611.1; JOINED.

DR EMBL; M62463; AAA50611.1; JOINED.

DR EMBL; M62464; AAA50611.1; JOINED.

DR EMBL; M62465; AAA50611.1; JOINED.

DR EMBL; M62466; AAA50611.1; JOINED.

DR EMBL; M62467; AAA50611.1; JOINED.

DR EMBL; M62468; AAA50611.1; JOINED.

DR EMBL; M62469; AAA50611.1; JOINED.

DR EMBL; M87276; AAA53063.1; -.

DR EMBL; J05606; AAA40431.1; -.

DR PIR; A40558; AAA40431.1; JOINED.

DR MGD; MGI:98737; Thbs1.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR006209; EGF_Like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR003367; tsp_3.

DR InterPro; IPR003129; TSPN.

DR InterPro; IPR001007; VWF_C.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00090; tsp_1; 3.

DR Pfam; PF02412; tsp_3; 8.

DR Pfam; PF02210; TSPN; 1.

DR Pfam; PF00093; vwc; 1.

DR SMART; SM00181; EGF; 3.

DR SMART; SM00209; TSP1; 3.

DR SMART; SM00210; TSPN; 1.

DR SMART; SM00214; VWC; 1.

DR PROSITE; PS00022; EGF_1; FALSE_NEG.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS50092; TSP1; 3.

DR PROSITE; PS01208; VWF_C_1; 1.

DR PROSITE; PS50184; VWF_C_2; 1.

KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;

KW EGF-like domain; Signal.

FT SIGNAL 1 18

FT CHAIN 19 1170 THROMBOSPONDIN 1.

FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).

FT DOMAIN 24 221 TSP N-TERMINAL.

FT DOMAIN 316 373 VWF_C.

FT DOMAIN 379 429 TSP TYPE-1 1.

FT DOMAIN 435 490 TSP TYPE-1 2.

FT DOMAIN 492 547 TSP TYPE-1 3.

FT DOMAIN 549 587 EGF-LIKE 1.

FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 646 690 EGF-LIKE 3.

FT DOMAIN 723 758 TSP TYPE-3 1.

FT DOMAIN 759 781 TSP TYPE-3 2.

FT DOMAIN 782 817 TSP TYPE-3 3.

FT DOMAIN 818 840 TSP TYPE-3 4.

FT DOMAIN 841 878 TSP TYPE-3 5.

FT DOMAIN 879 914 TSP TYPE-3 6.

FT DOMAIN 915 950 TSP TYPE-3 7.

FT DOMAIN 951 1170 C-TERMINAL.

FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).

FT DISULFID 270 270 INTERCHAIN (PROBABLE).

FT DISULFID 274 274 INTERCHAIN (PROBABLE).

FT DISULFID 391 423 BY SIMILARITY.

FT DISULFID 395 428 BY SIMILARITY.

FT DISULFID 406 413 BY SIMILARITY.

FT DISULFID 447 484 BY SIMILARITY.

FT DISULFID 451 489 BY SIMILARITY.

FT DISULFID 462 474 BY SIMILARITY.

FT DISULFID 504 541 BY SIMILARITY.

FT DISULFID 508 546 BY SIMILARITY.

FT DISULFID 519 531 BY SIMILARITY.

FT DISULFID 551 562 BY SIMILARITY.

FT DISULFID 556 572 BY SIMILARITY.

FT DISULFID 575 586 BY SIMILARITY.

FT DISULFID 592 608 BY SIMILARITY.

FT DISULFID 599 617 BY SIMILARITY.

FT DISULFID 620 644 BY SIMILARITY.

FT DISULFID 650 663 BY SIMILARITY.

FT DISULFID 657 676 BY SIMILARITY.

FT DISULFID 678 689 BY SIMILARITY.

FT DISULFID 705 713 BY SIMILARITY.

FT DISULFID 718 738 BY SIMILARITY.

FT DISULFID 754 774 BY SIMILARITY.

FT DISULFID 777 797 BY SIMILARITY.

FT DISULFID 813 833 BY SIMILARITY.

FT DISULFID 836 856 BY SIMILARITY.

FT DISULFID 874 894 BY SIMILARITY.

FT DISULFID 910 930 BY SIMILARITY.

FT DISULFID 946 1167 BY SIMILARITY.

FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 1025 1025 F -> L (IN REF. 2).

SQ SEQUENCE 1170 AA; 129646 MW; 0443E493615E7F06 CRC64;

Query Match 31.1%; Score 50; DB 1; Length 1170;

Best Local Similarity 40.7%; Pred. No. 74;

Matches 11; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

Oy 2 YVKVDSCEEPQLINKNEAEADYDDDD 28

Db 712 HCKKDCNPILP-----NSGQEDYDKD 732

RESULT 12

ITP2_MOUSE

ID ITP2_MOUSE STANDARD; PRT: 350 AA.

AC Q9R000; Q9D161;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin beta-1 binding protein 2 (Melusin).
 GN ITGB1BP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=99436136; PubMed=10506186;
 RA Braccaccio M., Guazzone S., Menini N., Sibona E., Hirsch E.,
 RA De Andrea M., Rocchi M., Altura F., Tarone G., Silengo L.,
 RA "Melusin is a new muscle-specific interactor for beta(1) integrin
 RT cytoplasmic domain".
 RL J. Biol. Chem. 274:29282-29288(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakada I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner I., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Gustincich S., Hill D., Fletcher C., Fujita M., Gariboldi M.,
 RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: MAY PLAY A ROLE DURING MATURATION AND/OR ORGANIZATION OF
 CC MUSCLES CELLS.
 CC -!- SUBUNIT: INTERACTS WITH BETA-1 INTEGRIN SUBUNIT. THIS INTERACTION
 CC IS REGULATED BY DIVALENT CATIONS, AND IT OCCURS ONLY IN ABSENCE OF
 CC CALCIUM.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLES BUT
 CC NOT IN OTHER TISSUES. IS LOCALIZED IN ROWS FLANKING THE Z LINE
 CC CONTAINING ALPHA-ACTININ.
 CC -!- DEVELOPMENTAL STAGE: DETECTABLE IN EMBRYO LIMBS AT DAY 15, REACHED
 CC A MAXIMUM IN NEWBORN, AND DECLINED IN ADULT LIMB MUSCLES.
 CC DURING HEART DEVELOPMENT LEVEL REMAINS STEADY FROM EMBRYONIC DAY
 CC 15 TO ADULT STAGE.
 CC -!- INDUCTION: DURING MUSCLE REGENERATION.
 CC -!- DOMAIN: THE TAIL DOMAIN BINDS TO THE CYTOPLASMIC DOMAIN OF BOTH
 CC INTEGRIN BETA-1A AND BETA-1D ISOFORMS. THE PRESENCE OF CA(2+) IONS
 CC DOES NOT PREVENT BINDING OF A FRAGMENT CONSISTING OF THE SECOND
 CC CYSTEINE RICH REPEAT AND THE TAIL DOMAIN BUT PREVENTS THE BINDING
 CC OF THE FULL-LENGTH PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF140691; AAT01677.1;
 CC EMBL: AK003906; BA523069.1;
 CC MGD: MGI:1353420; Itgb1bp2.
 CC InterPro: IPR007051; CHORD.

DR InterPro: IPR007052; CS.
 DR Pfam: PF04968; CHORD; 2.
 DR Pfam: PF04969; CS; 1.
 KW SH3-binding.
 FT DOMAIN 5 59 CYS-RICH.
 FT SITE 150 204 CYS-RICH.
 FT SITE 28 31 SH3-BINDING (POTENTIAL).
 FT SITE 70 79 SH3-BINDING (POTENTIAL).
 FT SITE 173 176 SH3-BINDING (POTENTIAL).
 FT SITE 159 162 SH2-BINDING (POTENTIAL).
 FT SITE 235 238 SH2-BINDING (POTENTIAL).
 FT DOMAIN 321 350 ASP/GLU-RICH (ACIDIC).
 FT CONFLICT 148 148 S -> F (IN REF. 2).
 FT CONFLICT 172 172 H -> Y (IN REF. 2).
 FT CONFLICT 182 189 MKSWSCG -> KKFLLKLC (IN REF. 2).
 SQ SEQUENCE 350 AA; 38767 MW; 6DEBCCF805DE9289 CRC64;
 Query Match 30.7%; Score 49.5; DB 1; Length 350;
 Best Local Similarity 42.9%; Pred. No. 23;
 Matches 12; Conservative 6; Mismatches 7; Indels 3; Gaps 1;
 QY 6 DSCPEPQ---LIMKNNEEAEDYDDLT 30
 Db 308 DSAEKARAGVLELMEDEESEDSDLS 335
 || | : | : | : |||||
 || | : | : | : |||||
 RESULT 13
 ID EF2_AERPE STANDARD; PRT; 736 AA.
 AC Q9YC19;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Elongation factor 2 (EF-2).
 GN FUSA OR FUS OR APEL432.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF/GEF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP000061; BAA80429.1;
 CC PIR: G72621; G72621.
 CC HSP: P13551; IFNM.
 CC HAMAP: MF_00054; -. 1.
 CC InterPro: IPR004543; aEF-2.
 CC InterPro: IPR000795; EF_GTPbind.

```

DR InterPro: IPR000640; EFG_C.
DR InterPro: IPR005517; EFG_IV.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00679; EFG_C; 1.
DR Pfam: PF03764; EFG_IV; 1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNCT.
DR TIGRfams: TIGR00490; aef-2; 1.
DR TIGRfams: TIGR00231; small_gtp; 1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 28 35 GTP (BY SIMILARITY).
FT NP_BIND 94 98 GTP (BY SIMILARITY).
FT NP_BIND 148 151 GTP (BY SIMILARITY).
FT MOD_RES 602 602 DIPHTHAMIDE (BY SIMILARITY).
SQ SEQUENCE 736 AA; 81960 MW; 43981490CB5E8D64 CRC64;

Query Match 30.7%; Score 49.5; DB 1; Length 736;
Best Local Similarity 32.4%; Pred. No. 52;
Matches 11; Conservative 5; Mismatches 7; Indels 11; Gaps 1;

QY 3 VKVDSCP-----EEPOLIMKNEEADY 25
||: | | | | | | | | | | | | | |
Db 403 VKIQDLPKIEALRLTIEDPNLVVVKINEETGEY 436

RESULT 14
GTR3_RABBIT
ID GTR3_RABBIT STANDARD; PRT; 400 AA.
AC Q9XSC2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3, brain) (Fragment).
GN SLC2A3 OR GLUT3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP STRAIN=New Zealand white;
RX MEDLINE=99196804; PubMed=10095018;
RA Devaskar S.U., Rajakumar P.A., Mink R.B., McKnight R.A.,
RA Thamocharan S., Hicks S.M.;
RT "Effect of development and hypoxia-ischemia upon rabbit brain glucose
RT transporter expression.";
RL Brain Res. 823:113-128(1999).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC GLUCOSE TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF117812; AAD26251.1; -.
CC InterPro: IPR007114; MFS.
CC InterPro: IPR005828; Sub.transporter.
CC InterPro: IPR005829; Sug.transporter.
CC Pfam: PF00083; sugar_tr; 1.
CC PRINTS: PR00171; SUGTRNSPORT.

```

```

DR TIGRfams: TIGR00879; SP; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transpor; Multigene family.
FT NON_TER 1 1
FT DOMAIN <1 4 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 5 25 3 (POTENTIAL).
FT DOMAIN 26 27 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 28 48 4 (POTENTIAL).
FT DOMAIN 49 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 5 (POTENTIAL).
FT DOMAIN 79 84 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 85 105 6 (POTENTIAL).
FT DOMAIN 106 174 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 175 195 7 (POTENTIAL).
FT DOMAIN 196 209 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 210 230 8 (POTENTIAL).
FT DOMAIN 231 238 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 239 259 9 (POTENTIAL).
FT DOMAIN 260 273 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 274 294 10 (POTENTIAL).
FT DOMAIN 295 304 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 305 325 11 (POTENTIAL).
FT DOMAIN 326 332 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 333 353 12 (POTENTIAL).
FT DOMAIN 354 400 CYTOPLASMIC (POTENTIAL).
FT SITE 182 184 DEFINES SUBSTRATE SPECIFICITY (BY
FT SIMILARITY).
SQ SEQUENCE 400 AA; 43722 MW; F845D9E5248B9A79 CRC64;

Query Match 30.4%; Score 49; DB 1; Length 400;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 CPEPQLIMKNNERAE 23
||| | : : | | |
Db 108 CPSPRFLINKKEDE 123

RESULT 15
VPS5_YEAST
ID VPS5_YEAST STANDARD; PRT; 675 AA.
AC Q92331; Q08483;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotated update)
DE Vacuolar protein sorting-associated protein VPS5.
DE VPS5 OR GRD2 OR YOR059W OR YOR29-20.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97318765; PubMed=9175702;
RA Nothwehr S.F., Hines A.E.;
RT "The yeast VPS5/GRD2 gene encodes a sorting nexin-1-like protein
RT required for localizing membrane proteins to the late Golgi.";
RL J. Cell Sci. 110:1063-1072(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97431786; PubMed=9285823;
RA Horadzovsky B.F., Davies B.A., Seaman M.N.J., McLaughlin S.A.,
RA Yoon S., Emr S.D.;
RT "A sorting nexin-1 homologue, Vps5p, forms a complex with Vps17p and
RT is required for recycling the vacuolar protein-sorting receptor.";
RL Mol. Biol. Cell 8:1529-1541(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97279235; PubMed=9133743;
RA Valens M., Bohn C., Baignan-Fornier B., Dang V., Bolotin-Fukuhara M.;
RT "The sequence of a 54.7 kb fragment of yeast chromosome XV reveals
RT the presence of two tRNAs and 24 new open reading frames.";

```

```
RL Yeast 13:379-390(1997).
CC -!- FUNCTION: REQUIRED FOR RETENTION OF LATE GOLGI MEMBRANE PROTEINS
CC AND VACUOLAR BIOGENESIS.
CC -!- SUBUNIT: INTERACTS WITH VPS17.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: PHOSPHORYLATED ON SERINE RESIDUE(S).
CC -!- SIMILARITY: BELONGS TO THE SORTING NEXIN FAMILY.
CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 13.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U73512; AAB62976.1; -.
CC EMBL; U84735; AAB41798.1; -.
CC EMBL; Z70678; CAA94554.1; ALT_FRAME.
CC EMBL; Z74977; CAA99262.1; -.
CC SGD; S0005595; VPS5.
CC InterPro: IPR001683; PX.
CC Pfam; PF00787; PX; 1.
CC SMART; SM00312; PX; 1.
CC PROSITE; PS0195; PX; 1.
CC Transport; Protein transport; Golgi stack; Phosphorylation.
CC DOMAIN 279 394
CC FT DOMAIN 279 394
CC SQ SEQUENCE 675 AA; 76484 MW; FBBA994EADBC2BD CRC64;

Query Match 29.8%; Score 48; DB 1; Length 675;
Best Local Similarity 36.4%; Pred. No. 75;
Matches 12; Conservative 7; Mismatches 8; Indels 6; Gaps 2;

QY 3 VKVDSQPE--EPQLMKNNEEADYD---DDL 29
DB 58 IDLSAPFWKDPGLSVAGNPQLEEDNSKADDL 90
:|||||:| |:|:|:| |
```

Search completed: October 21, 2003, 18:54:17
Job time : 11.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:42 ; Search time 44.75 Seconds
(without alignments)
172.996 Million cell updates/sec

Title: FVIII_ARG355I
Perfect score: 161
Sequence: 1 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	120.5	74.8	2343	6	O62730 canis famil
2	120.5	74.8	2343	6	O18806 canis famil
3	64	39.8	355	11	Q8BQ43 mus musculu
4	60.5	37.6	2454	5	Q8I3A6 plasmodium
5	57.5	35.7	192	5	Q9VQ17 drosophila
6	57	35.4	181	5	O8IIA0 plasmodium
7	57	35.4	1007	11	Q8C8T6 mus musculu
8	57	35.4	1755	11	Q8CHC7 mus musculu
9	56	34.8	1102	3	Q08964 saccharomyc
10	56	34.8	1402	4	Q9Y6X7 homo sapien
11	55.5	34.5	1085	3	Q8TGW9 candida alb
12	55.5	34.5	2447	10	Q94I07 oryza sativ
13	55	34.2	237	16	Q8D145 yersinia pe
14	55	34.2	330	2	Q8GJ97 yersinia ps
15	55	34.2	603	16	Q8FKV5 escherichia
16	53.5	33.2	390	10	Q9M2P2 arabidopsis

17	53	32.9	297	13	Q9DPB1	Q9dfb1 brachydanio
18	53	32.9	321	16	Q9ZUS6	Q9zls6 helicobacte
19	53	32.9	6632	5	Q01761	Q01761 caenorhabdi
20	53	32.9	6632	5	Q17362	Q17362 caenorhabdi
21	52	32.3	188	4	Q8N5I9	Q8n5i9 homo sapien
22	52	32.3	535	3	Q8J7X8	Q8j7x8 cryptococcu
23	52	32.3	537	2	Q85887	Q85887 sphingomona
24	52	32.3	649	4	Q96TD2	Q96td2 homo sapien
25	51.5	32.0	264	17	Q8TII5	Q8tii5 methanosarc
26	51.5	32.0	407	10	Q8I159	Q8i159 aethionema
27	51	31.7	428	10	Q9SJ54	Q9sjs4 arabidopsis
28	51	31.7	570	5	Q9T405	Q9u405 drosophila
29	51	31.7	938	10	Q9AV04	Q9av04 oryza sativ
30	51	31.7	977	16	Q8UF98	Q8ufq8 agrobacteri
31	51	31.7	1123	5	Q8IOW7	Q8iow7 plasmodium
32	51	31.7	1979	5	Q8IFX4	Q8icx4 plasmodium
33	51	31.7	2190	5	Q9VJMO	Q9vjm0 drosophila
34	50.5	31.4	134	6	Q9BEE2	Q9bee2 mustela lut
35	50.5	31.4	141	10	Q9M205	Q9m2q5 arabidopsis
36	50.5	31.4	612	11	Q8C056	Q8c0g6 mus musculu
37	50.5	31.4	622	10	Q945L6	Q945l6 arabidopsis
38	50.5	31.4	622	10	Q9CAW9	Q9caw9 arabidopsis
39	50.5	31.4	782	16	Q8EQS7	Q8eqs7 oceanobacil
40	50.5	31.4	1001	11	Q8BWH2	Q8bwh2 mus musculu
41	50.5	31.4	4638	5	Q8IK96	Q8ik96 plasmodium
42	50.5	31.4	6118	5	Q8I396	Q8i396 plasmodium
43	50	31.1	161	3	Q9Y709	Q9y709 schizosacch
44	50	31.1	170	16	Q8XSZ7	Q8xszt raiistonias
45	50	31.1	181	13	P79786	P79786 gallus gall

ALIGNMENTS

RESULT 1

O62730 PRELIMINARY; PRT; 2343 AA.
AC O62730;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Spleen;
RA Gordy P.W., Bowen R.A.;
RT "Characterization of the canine factor VIII cDNA."
RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL; AF049489; AAC05384.1; -.
DR HSSP; P00451; 1CFG.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASEL; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match

Best Local Similarity 74.8%; Score 120.5; DB 6; Length 2343;

Matches 24; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 AYVKVDSCEPPEQLIMKNEEAEDYDDDL 29

|||||

336 AYVKVDSCEPPEQLIMKNEED-KDYDDGL 363

Db 342 AYKVDSCPESQ 354

```

RESULT 4
GR13A6 PRELIMINARY; PRT; 2454 AA.
ID QBI3A6
AC QBI3A6;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10210C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RX MEDLINE=22255708; PubMed=12368667;
RA Hall N.K., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Lane A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares S., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL929355; CAD51728.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2454 AA; 292691 MW; 92018C7562CA79F0 CRC64;

Query Match 37.6%; Score 60.5; DB 5; Length 2454;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 14; Conservative 3; Mismatches 6; Indels 5; Gaps

QY 2 YKVDSCPPEPQLIMKNEEAEDYDDDL 29
Db 2309 YMKVDSITKE-----KNNNNDDDDDDL 2331

RESULT 5
Q9VQ17 PRELIMINARY; PRT; 192 AA.
ID Q9VQ17
AC Q9VQ17;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE CG18132 protein.
GN CG18132
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RX SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananaitides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.H., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA [1]

```

```

RL  Nature 419:498-511(2002).
DR  EMBL: AE014840; AAN35858.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 181 AA; 21718 MW; C4DDAAAD3701894D CRC64;

Query Match 35.4%; Score 57; DB 5; Length 181;
Best Local Similarity 47.6%; Pred. No. 4;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 5 VDSCPPEPQLMKNNEEAEDY 25
   : : : : :
Db 72 IDNIEKNIKNNEENY 92

RESULT 7
Q8C8T6 PRELIMINARY; PRT; 1007 AA.
ID Q8C8T6
AC Q8C8T6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Rho interacting protein 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
QC NCBI_TaxID=10090;
QX [1]
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK044512; BAC31958.1; -.
FT NON_TER 1007
SQ SEQUENCE 1007 AA; 1114976 MW; F49A2A0162190391 CRC64;

```

```

SQ      QUERY          1007 AA;   114976 MW;    F49A2A0162190391 CRG64;
Query Match           35.4%;   score 57;   DB 11;   Length 1007;
Best Local Similarity 37.9%;   pred.No. 23;
Matches 11; Conservative 8; Mismatches 6; Indels 4; Gaps 1;

QY      5 VDSCPPEPQLIMKKNNEEAED---YDDDL 29
       ::||::|::|::|::|::|::|::|::|
DB      871 LESCEQEQLLQNKEVEDKASAYEDQL 899

RESULT 8
Q8CHC7
ID      Q8CHC7        PRELIMINARY;      PRT; 1755 AA.
AC      Q8CHC7;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DR      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      MKTAAO864 protein (Fragment).
GN      MKTAAO864
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,
RA      Ohara O., Koga H.;
RT      "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT      I. The complete nucleotide sequences of 100 mouse KIAA-homologous
RT      cDNAs identified by screening of terminal sequences of cDNA clones
RT      randomly sampled from size-fractionated libraries.";
RL      Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; AB093269; BAC41453.1;
NR      NON TER 1

```

SO SEQUENCE 1755 AA; 198974 MW; 3591DC2DF9AA59FF CRC64;

Query Match 35.4%; Score 57; DB 11; Length 1755;
Best Local Similarity 37.9%; Pred. No. 41;
Matches 11; Conservative 8; Mismatches 6; Indels 4; Gaps 1;

QY 5 VDCPEEPQOLMKNEAEED----YDDDL 29

Db 357 LESCEKEKQALLQNLKEVEDKASAYEDQL 385

RESULT 9

Q08964 Q08964 PRELIMINARY; PRT; 1102 AA.

AC Q08964;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chromosome XVI reading frame ORF YPL216W.
GN YPL216W.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

[1]

SEQUENCE FROM N.A.

RA Rieger M., Mueller-Auer S., Schaefer M.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]

RA MIPS;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z73572; CAA97931.1; -.

DR SGD; S0006137; YPL216W.

DR InterPro; IPR001064; Crystallin.

DR InterPro; IPR004022; DDT_dom.

DR Pfam; PF02791; DDT_1.

DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.

SO SEQUENCE 1102 AA; 128132 MW; E457A8CE097255D6 CRC64;
Query Match 34.8%; Score 56; DB 3; Length 1102;
Best Local Similarity 45.8%; Pred. No. 35;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 6 DSCPPEPQOLMKNEAEEDYDDDL 29

Db 19 DSFETPWIKESSERINDYDSDL 42

RESULT 10

Q096X7 Q096X7 PRELIMINARY; PRT; 1402 AA.

AC Q096X7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE KIAA0864 protein (Fragment).

GN KIAA0864.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=99156230; PubMed=10048485;

RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosewa M.,

RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes.

RT XII.The complete sequences of 100 new cDNA clones from brain which

RT code for large proteins in vitro.";

RL DNA Res. 5:355-364(1998).

EMBL; AB020671; BAA74887.2; -.

FT NON_TER 1

SO SEQUENCE 1402 AA; 158411 MW; 4097CCBB25B3A179 CRC64;

Query Match 34.8%; Score 56; DB 4; Length 1402;
Best Local Similarity 37.9%; Pred. No. 45;
Matches 11; Conservative 8; Mismatches 6; Indels 4; Gaps 1;

QY 5 VDCPEEPQOLMKNEAEED----YDDDL 29

Db 3 LESCEKEKQALLQNLKEVEDKASAYEDQL 31

RESULT 11

Q08TGW9 Q08TGW9 PRELIMINARY; PRT; 1085 AA.

AC Q08TGW9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative transcriptional repressor.
GN SSN6.

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI_TaxID=5476;

[1]

SEQUENCE FROM N.A.

RA Hwang C.-S., Kang S.-O.;

RC STRAIN=ATCC10231;
RT "Molecular cloning of the gene encoding Candida albicans Prf1

RT homologous to Saccharomyces cerevisiae Ssn6.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF170083; AAL54912.2; -.

DR InterPro; IPR001440; TPR.

DR Pfam; PF00515; TPR; 9.

DR SMART; SM00028; TPR; 9.

SO SEQUENCE 1085 AA; 120524 MW; D38E0BAE275A0000 CRC64;

Query Match 34.5%; Score 55.5; DB 3; Length 1085;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 13; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 3 VKVDSCEPEPQOLMKNEAEEDYDDDL 28

Db 1063 VKKDENEPEP---MRKTEEDENYDDE 1085

RESULT 12

Q04107 Q04107 PRELIMINARY; PRT; 2447 AA.

AC Q04107;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative retroelement.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Echaritoidae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;

[1]

SEQUENCE FROM N.A.

RC STRAIN=Nipponbare;

RA McCombie W.R.;

RT "Rice genomic sequence.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=Nipponbare;

RA Cordum H., Johnson D., Minx P., de la Bastide M., Nascimento L.U.,

RA Spiegel L.A., Kirchoff K.A., King L., Preston R.R., Vil M.D.,

RA Baker J.P., Miller B., Santos L., Zutavern T., Kuit K.H.,

RA Rodriguez S., Cunnius D.M., Balija V.S., Shah R.S., Bahret A., Bal H.,

RC STRAIN=06:H1 / CFTQ73 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.I.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL: AE016755; AN78774.1; -
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 603 AA; 68354 MW; 13EFC0466FE11B6F CRC64;

Query Match 34.2%; Score 55; DB 16; Length 603;
 Best Local Similarity 35.7%; Pred. No. 26;
 Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 2 YVKVDSCEPEPQLIMKNEEAEDYDDL 29
 Db | | | | | : | | | : : : : :
 353 YRKRDPEPPEITLNNRREAIEYIEI 380

Search completed: October 21, 2003, 18:57:26
 Job time : 48.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:44:27 ; Search time 49.5 Seconds
(without alignments)
96.198 Million cell updates/sec

Title: FVIII_ARG581K

Perfect score: 153

Sequence: 1 LICYKESVDQKGNQIMSKRNVLFSVFDE 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	153	100.0	1383	18	AAW33228
2	153	100.0	2332	18	AAW33223
3	153	100.0	2332	18	AAW33224
4	153	100.0	2332	18	AAW33226
5	150	98.0	294	23	AAW79422
6	150	98.0	368	16	AAW73020
7	150	98.0	720	16	AAW74088
8	150	98.0	729	16	AAW74089
9	150	98.0	740	16	AAW76959

10	150	98.0	740	16	AAW76961	Human Factor-VIII
11	150	98.0	740	16	AAW76962	Human Factor-VIII
12	150	98.0	740	16	AAW73021	Human Factor-VIII
13	150	98.0	740	16	AAW74090	Factor-VIII heavy
14	150	98.0	740	16	AAW76982	Human factor VIII
15	150	98.0	1383	18	AAW33227	Procoagulant-activ
16	150	98.0	1383	18	AAW33229	Procoagulant-activ
17	150	98.0	1424	10	AAW91169	Sequence of 740 Ar
18	150	98.0	1424	22	AAW48842	Mutant mature huma
19	150	98.0	1424	23	AAW18622	Mutant mature B-dom
20	150	98.0	1425	9	AAW80267	Modified factor VI
21	150	98.0	1438	21	AAW01262	B-domain deleted f
22	150	98.0	1440	12	AAW12971	Factor VIII:SO. U
23	150	98.0	1445	23	ABG92540	LE B-domain-delete
24	150	98.0	1447	23	ABG92541	5Arg B-domain-dele
25	150	98.0	1457	19	AAW46246	Human factor VIII
26	150	98.0	1457	19	AAW44372	Human Factor VIII
27	150	98.0	1457	20	AAW21675	Beta-domain delete
28	150	98.0	1459	22	AAW10827	Human factor VIII
29	150	98.0	1459	22	AAW10832	Human factor VIII
30	150	98.0	1459	22	AAW10833	Human factor VIII
31	150	98.0	1471	18	AAW23414	Human B-domain del
32	150	98.0	1471	22	AAW67959	Amino acid sequenc
33	150	98.0	1516	9	AAW80265	Modified factor VI
34	150	98.0	1661	18	AAW18670	Factor VIII-dB695-
35	150	98.0	2098	17	AAW86863	Factor-VIII. Homo
36	150	98.0	2332	8	AAW71726	Factor VIII:C varia
37	150	98.0	2332	8	AAW71727	Factor VIII:C varia
38	150	98.0	2332	8	AAW71728	Factor VIII:C varia
39	150	98.0	2332	8	AAW71729	Factor VIII:C varia
40	150	98.0	2332	14	AAW43257	Human Factor VIII.
41	150	98.0	2332	18	AAW33222	Procoagulant-activ
42	150	98.0	2332	18	AAW33225	Procoagulant-activ
43	150	98.0	2332	19	AAW53483	Human factor VIII.
44	150	98.0	2332	19	AAW44132	Homo sapiens modif
45	150	98.0	2332	20	AAW31594	Human factor VIII

ALIGNMENTS

RESULT 1

AAW33228

ID AAW33228 standard; protein; 1383 AA.

XX AAW33228;

XX 30-APR-1998 (first entry)

DT Procoagulant-active human factor VIII:C (FVIII) mutant protein.

DE Procoagulant active factor VIII; FVIII: haemophilia A;

DE recombinant secretion; pro-coagulant activity; resistance;

KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;

KW von Willebrand factor binding site; binding affinity;

KW FVIII replacement therapy.

XX Synthetic.

OS Homo sapiens.

XX Key

XX Region

XX Location/Qualifiers

XX 1..346

XX /note= "factor VIIIA heavy chain"

XX 741..1383

XX /note= "factor VIIIA light chain"

XX 1..329

XX /note= "A1 domain"

XX 1..179

XX /note= "plastocyanin-like domain 1"

XX 187..329

XX /note= "plastocyanin-like domain 2"

XX 380..711

XX /note= "A2 domain"

FT Misc-difference 336 /note= "probable"
 FT /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 PN WO9740145-A1.
 XX 30-OCT-1997.
 XX 24-APR-1997; 97WO-US06563.
 XX 15-MAY-1996; 96US-0017785.
 PR 24-APR-1996; 96US-0016117.
 XX (UNMI) UNIV MICHIGAN.
 XX Amano K, Kaufman RJ, Pipe SW;
 XX WPI; 1997-535830/49.
 XX Modified human pro-coagulant active factor VIII - can be
 PT administered to haemophiliacs, i.e. factor VIII replacement therapy
 XX Claim 10; Page -: 57pp; English.
 XX The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising a mutation at Arg562, which is
 CC replaced with Lys. Factor VIII, along with calcium and phospholipid,
 CC acts as a cofactor for factor IXA, when it converts factor X to the
 CC activated form (factor IXa). FVIII is the coagulation factor deficient in
 CC the X-chromosome-linked bleeding disorder haemophilia A. Several other
 CC mutant FVIII proteins have also been created (see AAW33222-29). The
 CC FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
 CC higher levels than typically obtained with wild type FVIII and retains
 CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
 CC (AAW33223) are resistant to activated protein C (APC) cleavage. The
 CC FVIII mutant comprising a deletion of the B domain and von Willebrand
 CC factor binding site, a mutation at Arg740 and an addition of an amino
 CC acid sequence spacer between the A2 and A3 domains can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.
 XX SQ Sequence 2332 AA;
 Query Match 100.0%; Score 153; DB 18; Length 2332;
 ID Best Local Similarity 100.0%; Pred. No. 2e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
 Db |||||
 552 LICYKESVDQGNQIMSDKRNVLFSVFDE 581
 RESULT 3
 AAW33224
 ID AAW33224 standard; protein; 2332 AA.
 XX AC AAW33224;
 XX 30-APR-1998 (first entry)
 XX Procoagulant-active human factor VIII:C R336I/R562K mutant protein.
 DE Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;

KW FVIII replacement therapy.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Region 1..1313
 FT /note= "factor VIIIA heavy chain"
 FT Region 1649..2332
 FT /note= "factor VIIIA light chain"
 FT Domain 1..329
 FT /note= "A1 domain"
 FT Domain 1..179
 FT /note= "plastocyanin-like domain 1"
 FT Domain 187..329
 FT /note= "plastocyanin-like domain 2"
 FT Domain 380..711
 FT /note= "A2 domain"
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 741..1648
 FT /note= "B domain"
 FT Domain 1694..2021
 FT /note= "A3 domain"
 FT Domain 1694..1858
 FT /note= "plastocyanin-like domain 5"
 FT Domain 1888..2021
 FT /note= "plastocyanin-like domain 6"
 FT Domain 2021..2169
 FT /note= "C1 domain"
 FT Domain 2174..2326
 FT /note= "C2 domain"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Cleavage-site 740..741
 FT /note= "by thrombin"
 FT Cleavage-site 1689..1690
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Disulfide-bond 1832..1858
 FT /note= "probable"
 FT Misc-difference 336
 FT /label= R336I
 FT /note= "wild type Arg replaced with Ile"
 FT Misc-difference 562
 FT /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 XX WO9740145-A1.
 XX 30-OCT-1997.
 XX 24-APR-1997; 97WO-US06563.
 XX 15-MAY-1996; 96US-0017785.
 PR 24-APR-1996; 96US-0016117.
 XX (UNMI) UNIV MICHIGAN.
 XX Amano K, Kaufman RJ, Pipe SW;
 XX WPI; 1997-535830/49.
 XX Modified human pro-coagulant active factor VIII - can be
 PT administered to haemophiliacs, i.e. factor VIII replacement therapy
 XX Claim 10; Page -: 57pp; English.


```

XX SQ Sequence 2332 AA;
Query Match 100.0%; Score 153; DB 18; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 552 LICYKESVDQKGNQIMSDKRNVLFSVFDE 581

RESULT 5
AAG79422
ID AAG79422 standard; protein; 294 AA.
XX AC AAG79422;
XX DT 25-OCT-2002 (first entry)
XX DE Human factor VIII region 2/3.
XX KW Factor VIII; factor IX; coagulation; blood; cardiovascular disorder;
XX KW thrombosis; atherosclerosis; restenosis.
XX OS Homo sapiens.
XX PN W0200259268-A2.
XX PD 01-AUG-2002.
XX PF 23-JAN-2002; 2002WO-US01724.
XX PR 23-JAN-2001; 2001US-263431P.
XX PA (UYSL-) UNIV SAINT LOUIS.
XX PI Bajaj PS, Fay PJ;
XX DR WPI; 2002-599771/64.
XX PT New polypeptide inhibits interaction of blood coagulation factor VIIa
XX PT with factor IXa precluding activation of factor X; useful for
XX PT preventing or treating coagulation disorders, such as thrombosis,
XX PT atherosclerosis and restenosis -
XX PS Claim 2; Fig 7A; 61pp; English.
XX CC The sequences given in AAG79422-23 represent regions 2 and 3 of human
XX CC factors VIII and IX. Region 2 comprises the interaction between N346
XX CC of factor IXa and E455 and K570 of factor VIIa, and the interaction
XX CC between R403 of factor IXa and E633 of factor VIIa. Region 3 comprises
XX CC the interaction between K293 of factor IXa and D712 of factor VIIa, and
XX CC the interaction between E410 of factor IXa and K713 of factor VIIa.
XX CC Fragments of these polypeptides inhibit the interaction of blood
XX CC coagulation factor VIIa with blood coagulation factor IXa. They also
XX CC inhibit the activation of blood coagulation factor X, or inhibit blood
XX CC coagulation. These peptide fragments are useful in preventing and/or
XX CC treating coagulation disorders, in particular cardiovascular disorders
XX CC such as thrombosis, atherosclerosis and restenosis.
XX SQ Sequence 294 AA;
Query Match 98.0%; Score 150; DB 23; Length 294;
Best Local Similarity 96.7%; Pred. No. 5e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 123 LICYKESVDQKGNQIMSDKRNVLFSVFDE 152

RESULT 6

```

```

AAR73020
ID AAR73020 standard; peptide; 368 AA.
XX AC AAR73020;
XX DT 25-MAR-2003 (updated)
XX DT 21-NOV-1995 (first entry)
XX DE Human Factor-VIII fragment.
XX KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX OS Homo sapiens.
XX PN W09513301-A1.
XX PD 18-MAY-1995.
XX PF 10-NOV-1994; 94WO-DK00424.
XX PR 12-NOV-1993; 93DK-0001281.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Persson E;
XX WPI; 1995-194038/25.
XX PT Crosslinked Factor VIII polypeptide which is stable - is prep'd. using
XX PT bis(sulphosuccinimidy) suberate or disuccinimidy suberate in the
XX PT presence of polysorbate 80 to produce a coagulant with long lasting
XX PT activity
XX PS Disclosure; Page 19; 36pp; English.
XX CC This is a fragment corresponding to internal AAs 373-740 of human
XX CC Factor-VIII which may be crosslinked resulting in
XX CC increased stability and retention of high activity over extended
XX CC periods of time after activation by thrombin. The polypeptide is
XX CC used to prevent or treat diseases caused by the absence or deficiency
XX CC of Factor-VIII in a subject such as haemophilia.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 368 AA;
Query Match 98.0%; Score 150; DB 16; Length 368;
Best Local Similarity 96.7%; Pred. No. 6.5e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 180 LICYKESVDQKGNQIMSDKRNVLFSVFDE 209

RESULT 7
AAR74088
ID AAR74088 standard; protein; 720 AA.
XX AC AAR74088;
XX DT 25-MAR-2003 (updated)
XX DT 04-NOV-1995 (first entry)
XX DE Factor-VIII heavy chain N-terminal fragment.
XX KW human; Factor VIII; heavy chain; N-terminal fragment;
XX KW thrombin cleavage; blood-clotting.
XX OS Homo sapiens.
XX PN W09513300-A1.
XX PD 18-MAY-1995.

```

```

XX PF 10-NOV-1994; 94WO-DK00423.
XX PR 12-NOV-1993; 93DK-0001280.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Ezban Rasmussen M, Kjalke M;
XX PW 1995-194037/25.
XX
XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX Al-A2 domain - are easier to produce recombinantly and retain coagulant
XX activity, may be used to treat patients who have developed antibodies to
XX C-terminal epitope(s) of Factor VIII
XX
XX Claim 3; Page 24-26; 5lpp; English.
XX
XX The sequence represents N-terminal residues 1-720 of a human Factor-
XX VIII heavy chain. The sequence is shorter than the Al-A2 domain,
XX and is produced by treating a polypeptide containing the full Al-A2
XX domain of full-length Factor-VIII with a protease, e.g. thrombin.
XX The fragment has the same coagulant specific activity as full-length
XX Factor-VIII in a chromogenic assay, and is activated by thrombin at
XX a similar rate. The fragment may be produced recombinantly to reduce
XX production costs and improve safety, and production levels and
XX stability are higher than for the full-length form. The fragment may
XX be used to treat patients who have developed antibodies against
XX epitopes in the C-terminal part of the heavy chain.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Query Match 98.0%; Score 150; DB 16; Length 720;
XX Best Local Similarity 96.7%; Pred. No. 1.5e-14;
XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LICYESVDQGNQIMSDKRNVLFSVFDE 30
XX |||||:|||||:|||||:|||||:|||||
XX Db 552 LICYESVDQGNQIMSDKRNVLFSVFDE 581
XX
XX RESULT 8
XX AAR74089
XX ID AAR74089 standard; protein; 729 AA.
XX AC AAR74089;
XX
XX DT 25-MAR-2003 (updated)
XX DT 04-NOV-1995 (first entry)
XX
XX Factor-VIII heavy chain N-terminal fragment.
XX human; Factor VIII; heavy chain: N-terminal fragment;
XX thrombin cleavage; blood-clotting.
XX
XX Homo sapiens.
XX
XX W09513300-A1.
XX
XX PD 18-MAY-1995.
XX
XX PF 10-NOV-1994; 94WO-DK00423.
XX PR 12-NOV-1993; 93DK-0001280.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Ezban Rasmussen M, Kjalke M;
XX PW 1995-194037/25.
XX

```

```

XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX Al-A2 domain - are easier to produce recombinantly and retain coagulant
XX activity, may be used to treat patients who have developed antibodies to
XX C-terminal epitope(s) of Factor VIII
XX
XX Claim 2; Page 27-29; 5lpp; English.
XX
XX The sequence represents N-terminal residues 1-729 of a human Factor-
XX VIII heavy chain. The sequence is shorter than the Al-A2 domain,
XX and is produced by treating a polypeptide containing the full Al-A2
XX domain of full-length Factor-VIII with a protease, e.g. thrombin.
XX The fragment has the same coagulant specific activity as full-length
XX Factor-VIII in a chromogenic assay, and is activated by thrombin at
XX a similar rate. The fragment may be produced recombinantly to reduce
XX production costs and improve safety, and production levels and
XX stability are higher than for the full-length form. The fragment may
XX be used to treat patients who have developed antibodies against
XX epitopes in the C-terminal part of the heavy chain.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 729 AA;
XX
XX Query Match 98.0%; Score 150; DB 16; Length 729;
XX Best Local Similarity 96.7%; Pred. No. 1.5e-14;
XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LICYESVDQGNQIMSDKRNVLFSVFDE 30
XX |||||:|||||:|||||:|||||:|||||
XX Db 552 LICYESVDQGNQIMSDKRNVLFSVFDE 581
XX
XX RESULT 9
XX AAR76959
XX ID AAR76959 standard; protein; 740 AA.
XX AC AAR76959;
XX
XX DT 25-MAR-2003 (updated)
XX DT 09-MAR-1996 (first entry)
XX
XX Human Factor-VIII/Factor-VIIIa derivative.
XX
XX Factor-VIII; therapeutic; blood-clotting.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 403 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 404 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 433 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 482 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 500 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 434 /label= Asp absent or Asn, Ser, Thr, Ala
XX FT Misc-difference 440 /label= Glu absent or Gln, Ser, Thr, Ala
XX FT Misc-difference 440 /label= Glu absent or Gln, Ser, Thr, Ala
XX
XX W09518829-A1.
XX
XX PD 13-JUL-1995.
XX
XX PF 06-JAN-1995; 95WO-DK00010.
XX PR 07-JAN-1994; 94DK-0000030.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX

```

PI Ezban Rasmussen M, Nicolaisen EM, Persson E;
 XX WPI; 1995-255041/33.
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX Disclosure; Page 11-14; 24pp; English.
 XX The new Factor-VIII/Factor-VIIIa derivative comprises a functional
 CC A2 domain in which amino acid residues, as indicated in the
 CC features, are deleted or substituted by another amino acid so as to
 CC increase the overall positive charge. Asp-403, Asp-404, Asp-433,
 CC Asp-482 and Asp-500 are preferably substituted by Asn; Glu-434 and
 CC Glu-440 are preferably substituted by Gln. The new derivative has
 CC the same activity as the wild-type Factor-VIII but with improved
 CC stability (the activity is maintained for a longer period compared
 CC to the rapid decline of the activity of wt Factor-VIII). The new
 CC derivative can be used in a composition for treating diseases caused
 CC by an absence or deficiency of Factor-VIII, especially haemophilia.
 CC (Updated on 25-MAR-2003 to correct DR field.)
 XX Sequence 740 AA;
 SQ Query Match 98.0%; Score 150; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 1.5e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LICYKESVDQKGNOIMSDKRNVLFSVFDE 30
 DB |||||:|||||:|||||:|||||:|||||:|||||
 552 LICYKESVDQKGNOIMSDKRNVLFSVFDE 581
 RESULT 10
 AAR76961
 ID AAR76961 standard; protein; 740 AA.
 AC AAR76961;
 XX 09-MAR-1996 (first entry)
 DT Human Factor-VIII derivative.
 XX Factor-VIII; therapeutic; blood-clotting.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 692
 FT /label= absent or Ala, Thr, Ser, Gly or Asp
 FT Misc-difference 720
 FT /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729
 FT /label= absent or Val, Ala or Ile
 XX W09518827-A1.
 PN 13-JUL-1995.
 PD 06-JAN-1995; 95WO-DK000008.
 PF 07-JAN-1994; 94DK-0000032.
 PR (NOVO) NOVO-NORDISK AS.
 XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
 XX WPI; 1995-255039/33.
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.

XX Disclosure; Page 14-17; 30pp; English.
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is deleted or replaced with another amino acid
 CC residue, preferably Ser (see AAR76962). Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX Sequence 740 AA;
 SQ Query Match 98.0%; Score 150; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 1.5e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LICYKESVDQKGNOIMSDKRNVLFSVFDE 30
 DB |||||:|||||:|||||:|||||:|||||:|||||
 552 LICYKESVDQKGNOIMSDKRNVLFSVFDE 581
 RESULT 11
 AAR76962
 ID AAR76962 standard; protein; 740 AA.
 AC AAR76962;
 XX 09-MAR-1996 (first entry)
 DT Human Factor-VIII derivative.
 XX Factor-VIII; therapeutic; blood-clotting.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 692
 FT /label= Cys substituted by Ser
 FT Misc-difference 720
 FT /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729
 FT /label= absent or Val, Ala or Ile
 XX W09518827-A1.
 PN 13-JUL-1995.
 PD 06-JAN-1995; 95WO-DK000008.
 PF 07-JAN-1994; 94DK-0000032.
 PR (NOVO) NOVO-NORDISK AS.
 XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
 XX WPI; 1995-255039/33.
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX Disclosure; Page 18-20; 30pp; English.
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is replaced with Ser. For other (less preferred)
 CC substitutions at this site, see AAR76961. Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is

CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.

XX SQ Sequence 740 AA;

Query Match 98.0%; Score 150; DB 16; Length 740;

Best Local Similarity 96.7%; Pred. No. 1.5e-14;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSDKRNILFSVFDE 30

|||||:|||||:|||||:|||||:|||||
 Db 552 LICYKESVDQGNQIMSDKRNILFSVFDE 581

RESULT 12

AAR73021

ID AAR73021 standard; peptide; 740 AA.

XX

AC AAR73021;

XX

25-MAR-2003 (updated)

DT 21-NOV-1995 (first entry)

XX

Human Factor-VIII N-terminal fragment.

DE

Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.

KW

Homo sapiens.

OS

WO9513301-A1.

XX

18-MAY-1995.

PD

XX

10-NOV-1994; 94WO-DK00424.

PF

XX

12-NOV-1993; 93DK-0001281.

PR

XX

(NOVO) NOVO-NORDISK AS.

PA

XX

Persson E;

PI

XX

WPI; 1995-194038/25.

DR

XX

Crosslinked Factor VIII polypeptide which is stable - is prepd. using

bis(sulphosuccinimidy) suberate or disuccinimidy suberate in the

presence of polysorbate 80 to produce a coagulant with long lasting

activity

PT

XX

Disclosure; Page 21; 36pp; English.

PS

XX

This is the N-terminal fragment of human Factor-VIII which may be

crosslinked resulting in increased stability and retention of high

activity over extended periods of time after activation by thrombin.

The polypeptide is used to prevent or treat diseases caused by the

absence or deficiency of Factor-VIII in a subject such as

haemophilia.

(Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 740 AA;

Query Match 98.0%; Score 150; DB 16; Length 740;

Best Local Similarity 96.7%; Pred. No. 1.5e-14;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSDKRNILFSVFDE 30

|||||:|||||:|||||:|||||:|||||
 Db 552 LICYKESVDQGNQIMSDKRNILFSVFDE 581

RESULT 13

AAR74090

ID AAR74090 standard; protein; 740 AA.

XX

AC AAR74090;

XX

25-MAR-2003 (updated)

DT

XX

04-NOV-1995 (first entry)

DE

XX

Factor-VIII heavy chain N-terminal fragment.

XX

human; Factor VIII; heavy chain; N-terminal fragment;

KW

thrombin cleavage; blood-clotting.

XX

Homo sapiens.

OS

WO9513300-A1.

PN

XX

18-MAY-1995.

PD

XX

10-NOV-1994; 94WO-DK00423.

PF

XX

12-NOV-1993; 93DK-0001280.

PR

XX

(NOVO) NOVO-NORDISK AS.

PA

XX

Ezban Rasmussen M, Kjalke M;

PI

XX

WPI; 1995-194037/25.

DR

XX

Factor VIII polypeptide(s) comprising a heavy chain shorter than native

AI-A2 domain - are easier to produce recombinantly and retain coagulant

activity, may be used to treat patients who have developed antibodies to

C-terminal epitope(s) of Factor VIII

XX

Disclosure; Page 30-32; 51pp; English.

PS

XX

The sequence represents N-terminal residues 1-740 of a human Factor-

VIII heavy chain. The sequence contains entire A1 and A2 domains,

and truncated forms (1-720 in (AAR74088) and 1-729 in (AAR74090))

may be produced by treatment with a protease, e.g. thrombin. The

C-terminally truncated fragments have the same coagulant specific

activity as full-length Factor-VIII, and may be produced

recombinantly to reduce production costs and improve safety, giving

higher production levels and stability than for the full-length form.

The fragments may be used to treat patients who have developed

antibodies against epitopes in the C-terminal part of the heavy chain.

XX

(Updated on 25-MAR-2003 to correct PN field.)

CC

XX

SQ Sequence 740 AA;

Query Match 98.0%; Score 150; DB 16; Length 740;

Best Local Similarity 96.7%; Pred. No. 1.5e-14;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSDKRNILFSVFDE 30

|||||:|||||:|||||:|||||:|||||
 Db 552 LICYKESVDQGNQIMSDKRNILFSVFDE 581

RESULT 14

AAR76982

ID AAR76982 standard; protein; 740 AA.

XX

AC AAR76982;

XX

21-FEB-1996 (first entry)

DT

XX

Human factor VIII A2-domain derivative.

DE

XX

Factor VIII; human; haemophilia; thrombin; protein C; plasmin;

KW

serine protease; recombination; therapy; deficiency.

XX

Homo sapiens.

OS

XX

```

FH Key Location/Qualifiers
FT Misc-difference 720
FT /label= Gln, Ser, Thr, Val, Ala
FT Misc-difference 729
FT /label= Val, Ala, Ile
PN WO9518828-A1.
XX 13-JUL-1995.
XX
XX
XX 06-JAN-1995; 95WO-DK00009.
XX
XX 07-JAN-1994; 94DK-0000031.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Erban Rasmussen M, Nicolaisen EM, Persson E;
XX WPI; 1995-255040/33.
XX
XX Novel factor VIII derivative with resistance to enzymatic cleavage -
XX and comprises a functional A2 domain where Glu720 and/or Tyr729 is
XX deleted/substituted.
XX
XX Claim 1; Page 11-14; 26pp; English.
XX
XX This sequence represents the A2 domain of a human factor VIII derivative.
XX Factor VIII is a large glycoprotein which is present in plasma at low
XX concentrations. Factor VIII is an essential part of the clotting
XX reaction in response to a wound. Factor VIII is susceptible to cleavage
XX by thrombin, activated protein C, plasmin, and other serine proteases.
XX Full length factor VIII consists of three repeats of the A-domain, a
XX B-domain and 2 repeats of the C-domain. Active factor VIII has the A1
XX domain cleaved off. Factor VIII is too unstable for use in recombinant
XX techniques. Factor VIII containing this sequence has improved stability
XX and shows resistance against enzymatic activity present in mammalian
XX cells. This means that factor VIII containing this sequence can be used
XX in recombinant techniques. A factor VIII derivative can be used for
XX treating diseases caused by an absence or deficiency of factor VIII (in
XX the same way as normal factor VIII) e.g. haemophilia. The advantage with
XX using a recombinant factor VIII also includes no need for lots of donors
XX in order to get a sufficient amount. Also, there is no long purification
XX process, and there is no risk of transmission of blood-borne diseases
XX such as HIV.
XX
XX Sequence 740 AA;
XX
XX Query Match 98.0%; Score 150; DB 16; Length 740;
XX Best Local Similarity 96.7%; Pred. No. 1.5e-14;
XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFE 30
XX |||||:|||||:|||||:|||||:|||||
XX Db 552 LICYKESVDQKGNQIMSDKRNVLFSVFE 581
XX
XX RESULT 15
XX AAW33227
XX ID AAW33227 standard; protein; 1383 AA.
XX
XX AC AAW33227;
XX
XX 30-APR-1998 (first entry)
XX
XX Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
XX Pro-coagulant active factor VIII; FVIII; haemophilia A;
XX recombinant secretion; pro-coagulant activity; resistance;
XX activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
XX von Willebrand factor binding site; binding affinity;
XX FVIII replacement therapy.
XX
XX Synthetic.
XX
OS Homo sapiens.
XX Key Location/Qualifiers
XX 1..346
XX /note= "factor VIIIA heavy chain"
XX 741..1383
XX /note= "factor VIIIA light chain"
XX 1..329
XX /note= "A1 domain"
XX 1..179
XX /note= "plastocyanin-like domain 1"
XX 187..329
XX /note= "plastocyanin-like domain 2"
XX 380..711
XX /note= "A2 domain"
XX 711..746
XX /note= "a spacer of the sequence
XX SFSONSRHPSTROKQFNATIPENDIEKTDWPF
XX AHRTMPKIQNVSSDLLMLL is inserted
XX between domains A2 and A3"
XX
XX 380..554
XX /note= "plastocyanin-like domain 3"
XX 564..711
XX /note= "plastocyanin-like domain 4"
XX 746..1073
XX /note= "A3 domain"
XX 1073..1221
XX /note= "C1 domain"
XX 1226..1378
XX /note= "C2 domain"
XX 372..373
XX /note= "by thrombin"
XX 153..179
XX /note= "probable"
XX 528..554
XX /note= "probable"
XX 740
XX /label= R740A
XX /note= "wild type Arg replaced with Ala"
XX
XX WO9740145-A1.
XX 30-OCT-1997.
XX
XX 24-APR-1997; 97WO-US06563.
XX
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96US-0016117.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Amano K, Kaufman RJ, Pipe SW;
XX WPI; 1997-535830/49.
XX
XX Modified human pro-coagulant active factor VIII - can be
XX administered to haemophiliacs, i.e. factor VIII replacement therapy
XX
XX Claim 20; Page -: 57pp; English.
XX
XX The present sequence represents a novel pro-coagulant active factor
XX VIII (FVIII) mutant protein, comprising a deletion of the B domain and
XX von Willebrand factor binding site, a mutation at Arg740 and an addition
XX of an amino acid sequence spacer between the A2 and A3 domains. Factor
XX VII, along with calcium and phospholipid, acts as a cofactor for factor
XX IXA, when it converts factor X to the activated form (XA). FVIII
XX is the coagulation factor deficient in the X-chromosome-linked bleeding
XX disorder haemophilia A. Several other mutant FVIII proteins have also
XX been created (see AAW3322-29). The FVIII mutant F309S (AAW33225) is
XX capable of recombinant secretion at higher levels than typically obtained
XX with wild type FVIII and retains pro-coagulant activity. The FVIII
XX mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
XX protein C (APC) cleavage. The present FVIII mutant can form a more

```

CC stable configuration, and have an approximate 5-fold increase in
CC specific activity compared to purified wild type FVIII, while increasing
CC their binding affinity to von Willebrand factor improves their
CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
CC FVIII replacement therapy, while the nucleic acid molecule can be used
CC for gene therapy.
CC note: this sequence does not appear in the specification; it was created
CC using sequences from the given references.
XX

SQ Sequence 1383 AA;

Query Match 98.0%; Score 150; DB 18; Length 1383;
Best Local Similarity 96.7%; Pred. No. 3.2e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 30
|||||:|||||
Db 552 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 581

Search completed: October 21, 2003, 18:53:31
Job time : 50.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:50:07 ; Search time 16 Seconds
(without alignments)
79.333 Million cell updates/sec

Title: FVIII_ARG581K

Perfect score: 153

Sequence: 1 LCYKESVDQKGNQIMSDKRNVLFSVFDE 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	98.0	1438	4	US-09-209-916-1
2	150	98.0	1471	1	US-08-683-839B-3
3	150	98.0	1661	2	US-08-882-083-2
4	150	98.0	1661	2	US-08-558-107-2
5	150	98.0	1661	3	US-09-243-539-2
6	150	98.0	2332	1	US-07-864-004B-4
7	150	98.0	2332	1	US-08-251-937A-4
8	150	98.0	2332	1	US-08-212-133A-2
9	150	98.0	2332	1	US-08-276-594A-2
10	150	98.0	2332	1	US-08-474-503-2
11	150	98.0	2332	2	US-08-670-707A-2
12	150	98.0	2332	3	US-09-037-601-2
13	150	98.0	2332	3	US-09-324-867-3
14	150	98.0	2332	4	US-09-315-179-2
15	150	98.0	2332	4	US-09-523-656-2
16	150	98.0	2332	5	PCT-US93-03275-4
17	150	98.0	2332	5	PCT-US94-13200-2
18	150	98.0	2351	1	US-08-121-202-2
19	150	98.0	2351	1	US-08-366-851A-2
20	150	98.0	2351	6	5171844-2
21	150	98.0	2351	6	5422260-1
22	147	96.1	367	1	US-07-864-004B-2
23	147	96.1	367	1	US-08-251-937A-2
24	147	96.1	367	5	PCT-US93-03275-2
25	147	96.1	368	1	US-08-212-133A-6
26	147	96.1	368	1	US-08-474-503-4
27	147	96.1	368	2	US-08-670-707A-4

28 147 96.1 368 3 US-09-037-601-4 Sequence 4, Appli
29 147 96.1 368 4 US-09-315-179-4 Sequence 4, Appli
30 147 96.1 368 4 US-09-523-656-4 Sequence 4, Appli
31 147 96.1 368 5 PCT-US94-13200-4 Sequence 4, Appli
32 147 96.1 541 1 US-08-121-202-4 Sequence 4, Appli
33 147 96.1 1443 2 US-08-670-707A-39 Sequence 39, Appli
34 147 96.1 1443 3 US-09-037-601-39 Sequence 39, Appli
35 147 96.1 1443 4 US-09-315-179-39 Sequence 39, Appli
36 147 96.1 1467 4 US-09-523-656-38 Sequence 38, Appli
37 147 96.1 2133 2 US-08-670-707A-37 Sequence 37, Appli
38 147 96.1 2133 3 US-09-037-601-37 Sequence 37, Appli
39 147 96.1 2133 4 US-09-315-179-37 Sequence 37, Appli
40 147 96.1 2133 4 US-09-523-656-30 Sequence 30, Appli
41 147 96.1 2343 3 US-09-324-867-2 Sequence 2, Appli
42 146 95.4 2304 3 US-09-324-867-4 Sequence 4, Appli
43 146 95.4 2319 1 US-08-212-133A-8 Sequence 8, Appli
44 146 95.4 2319 1 US-08-474-503-6 Sequence 6, Appli
45 146 95.4 2319 2 US-08-670-707A-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-209-916-1

; Sequence 1, Application US/09209916

; Patent No. 6358703

; GENERAL INFORMATION:

; APPLICANT: Cho, Myung-Sam

; APPLICANT: Cho, Sham-Yuen

; APPLICANT: Kelsey, William

; APPLICANT: Yee, Helena

; TITLE OF INVENTION: Expression System for Factor VIII

; FILE REFERENCE: MSB-7255

; CURRENT APPLICATION NUMBER: US/09/209,916

; CURRENT FILING DATE: 1998-12-10

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1438

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Derived from

; OTHER INFORMATION: human factor VIII sequence

US-09-209-916-1

Query Match

Best Local Similarity 98.0%; Score 150; DB 4; Length 1438;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LCYKESVDQKGNQIMSDKRNVLFSVFDE 30

|||||||:|||||||

DB 552 LCYKESVDQKGNQIMSDKRNVLFSVFDE 581

RESULT 2

US-08-683-839B-3

; Sequence 3, Application US/08683839B

; Patent No. 5744326

; GENERAL INFORMATION:

; APPLICANT: Ill, Charles . R. et al.

; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional

; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of

; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street, suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-839B-3

```

```

Query Match          98.0%; Score 150; DB 1; Length 1471;
Best Local Similarity 96.7%; Pred. No. 3.6e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

```

```

RESULT 3
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-882-083-2
; Query Match          98.0%; Score 150; DB 2; Length 1661;
; Best Local Similarity 96.7%; Pred. No. 4.2e-15;
; Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

```

```

RESULT 4
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-107-2

```

```

Query Match          98.0%; Score 150; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 4.2e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

```

```

RESULT 5
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Poley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-243-539-2

```

```

Query Match          98.0%; Score 150; DB 3; Length 1661;
Best Local Similarity 96.7%; Pred. No. 4.2e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNQIMSDKRNILFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 571 LICYKESVDQKGNQIMSDKRNILFSVFDE 600

```

```

RESULT 6
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4

```

```

Query Match          98.0%; Score 150; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 5.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNQIMSDKRNILFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICYKESVDQKGNQIMSDKRNILFSVFDE 581

```

```

RESULT 7
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien

```

TISSUE TYPE: Liver cdna sequence
US-08-251-937A-4

Query Match 98.0%; Score 150; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICKESVDQKGQIMSKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICKESVDQKGQIMSKRNVLFSVFDE 581

RESULT 8

US-08-212-133A-2
Sequence 2, Application US/08212133A
Patent No. 5663060

GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cdna sequence

US-08-212-133A-2

Query Match 98.0%; Score 150; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICKESVDQKGQIMSKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICKESVDQKGQIMSKRNVLFSVFDE 581

RESULT 9

US-08-276-594A-2

TISSUE TYPE: Liver cdna sequence
US-08-251-937A-4

Query Match 98.0%; Score 150; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICKESVDQKGQIMSKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICKESVDQKGQIMSKRNVLFSVFDE 581

Sequence 2, Application US/08276594A
Patent No. 5693499
GENERAL INFORMATION:
APPLICANT: YONEMURA, Hiroshi
APPLICANT: TAJIMA, Yoshitaka
APPLICANT: SUGAWARA, Keishin
APPLICANT: MASUDA, Kenichi
TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,594A
FILING DATE: 18-JUL-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/950,191
FILING DATE: 24-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 243262/1991
FILING DATE: 24-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258

REFERENCE/DOCKET NUMBER: 74129/195/AOPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-276-594A-2

Query Match 98.0%; Score 150; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LICKESVDQKGQIMSKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICKESVDQKGQIMSKRNVLFSVFDE 581

RESULT 10

US-08-474-503-2

Sequence 2, Application US/08474503
Patent No. 5744446
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309

COMPUTER READABLE FORM:


```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

```

```

Query Match          98.0%; Score 150; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 581

```

RESULT 13

```

US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:

```

```

; APPLICANT: Lillcrap, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632ley, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.0010002/JAG/RJD
; CURRENT APPLICATION NUMBER: US/09/324,867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-059
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-324-867-3

```

```

Query Match          98.0%; Score 150; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 553 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 582

```

RESULT 14

```

US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:

```

```

; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200

```

```

; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-179-2

```

```

Query Match          98.0%; Score 150; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 581

```

RESULT 15

```

US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:

```

```

; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

```

```

Query Match          98.0%; Score 150; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 6.4e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 552 LICYKESVDQKGNGQIMSDKRNVLFSVFDE 581

```

```

Search completed: October 21, 2003, 18:59:49
Job time : 17 secs

```



```

; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match          98.0%; Score 150; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 1.4e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 552 LICYKESVDQGNQIMSDKRNVLFSVFDE 581

RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US2003007752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

Query Match          98.0%; Score 150; DB 15; Length 1438;
Best Local Similarity 96.7%; Pred. No. 1.4e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 552 LICYKESVDQGNQIMSDKRNVLFSVFDE 581

RESULT 4
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony

```

```

; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

```

```

Query Match          98.0%; Score 150; DB 14; Length 1471;
Best Local Similarity 96.7%; Pred. No. 1.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

```

```

RESULT 5
US-09-957-641-2
; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00
; CURRENT APPLICATION NUMBER: US/09/957,641
; CURRENT FILING DATE: 2001-09-16
; PRIOR APPLICATION NUMBER: US 60/234047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-957-641-2

```

```

Query Match          98.0%; Score 150; DB 10; Length 2332;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
Db 552 LICYKESVDQGNQIMSDKRNVLFSVFDE 581

```

```

RESULT 6
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601

```

```

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-510A-2

```

```

Query Match          98.0%; Score 150; DB 12; Length 2332;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 LICYESVDQKGQIMSDKRNVLFSVFDE 30
        |||||:|||||:|||||:|||||:|||||
Db      552 LICYESVDQKGQIMSDKRNVLFSVFDE 581

```

RESULT 7

```

US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: <Unknown>
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

```

```

Query Match          98.0%; Score 150; DB 15; Length 2332;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 LICYESVDQKGQIMSDKRNVLFSVFDE 30
        |||||:|||||:|||||:|||||:|||||
Db      552 LICYESVDQKGQIMSDKRNVLFSVFDE 581

```

RESULT 8

```

US-10-133-907-4
; Sequence 4, Application US/10133907
; Publication No. US2003019523A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R.
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pall70
; CURRENT APPLICATION NUMBER: US/10/133,907
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-907-4

```

```

Query Match          98.0%; Score 150; DB 12; Length 2351;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 LICYESVDQKGQIMSDKRNVLFSVFDE 30
        |||||:|||||:|||||:|||||:|||||
Db      571 LICYESVDQKGQIMSDKRNVLFSVFDE 600

```

RESULT 9

```

US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R.
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pall70
; CURRENT APPLICATION NUMBER: US/10/132,829
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-4

```


Query Match 98.0%; Score 150; DB 15; Length 2351;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQIMSDKRNILFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 571 LICYKESVDQKGQIMSDKRNILFSVFDE 600

RESULT 10
US-10-172-712-27
; Sequence 27, Application US/10172712
; Publication No. US20030125232A1
; GENERAL INFORMATION:
; APPLICANT: GRIFFIN, JOHN H.
; APPLICANT: GALE, ANDREW J.
; APPLICANT: GETZOFF, ELIZABETH D.
; APPLICANT: PELLEQUER, JEAN-LUC
; TITLE OF INVENTION: STABILIZED PROTEINS WITH ENGINEERED DISULFIDE BONDS
; FILE REFERENCE: 4198-4001US1
; CURRENT APPLICATION NUMBER: US/10/172,712
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 60/298,578
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-172-712-27

Query Match 98.0%; Score 150; DB 15; Length 2351;
Best Local Similarity 96.7%; Pred. No. 2.5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQIMSDKRNILFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 571 LICYKESVDQKGQIMSDKRNILFSVFDE 600

RESULT 11
US-10-131-510A-4
; Sequence 4, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-131-510A-4

Query Match 96.1%; Score 147; DB 12; Length 368;
Best Local Similarity 93.3%; Pred. No. 8e-14;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQIMSDKRNILFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 180 LICYKESVDQKGQIMSDKRNILFSVFDE 209

RESULT 12
US-10-187-319-4
; Sequence 4, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Porcine
; TISSUE TYPE: spleen
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..368
; OTHER INFORMATION:
; sequence of porcine factor VIII A2 domain, defined as residues
; homologous to human factor VIII, amino acids 373-740. Residues
; 1-4 are from known porcine amino acid sequence.
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-187-319-4

Query Match 96.1%; Score 147; DB 15; Length 368;

```
Best Local Similarity 93.3%; Pred. No. 8e-14;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQNMDSKRNVLFSVFE 30
    |||||:||||:||||:||||:||||:||||:
Db 180 LICYKESVDQKGQNMDSKRNVLFSVFE 209

RESULT 13
US-10-095-718-4
; Sequence 4, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: canine B-domain deleted factor VIII
US-10-095-718-4

Query Match 96.1%; Score 147; DB 14; Length 1431;
Best Local Similarity 93.3%; Pred. No. 4e-13;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQNMDSKRNVLFSVFE 30
    |||||:||||:||||:||||:||||:||||:
Db 565 LICYKESVDQKGQNMDSKRNVLFSVFE 594

RESULT 14
US-10-131-510A-39
; Sequence 39, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-957
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1443

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: factor VIII lacking
; OTHER INFORMATION: the B domain
US-10-131-510A-39

Query Match 96.1%; Score 147; DB 12; Length 1443;
Best Local Similarity 93.3%; Pred. No. 4e-13;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQNMDSKRNVLFSVFE 30
    |||||:||||:||||:||||:||||:||||:
Db 571 LICYKESVDQKGQNMDSKRNVLFSVFE 600

RESULT 15
US-10-187-319-39
; Sequence 39, Application US/10187319
; Publication No. US20030066785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-10-187-319-39

Query Match 96.1%; Score 147; DB 15; Length 1443;
Best Local Similarity 93.3%; Pred. No. 4e-13;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGQNMDSKRNVLFSVFE 30
    |||||:||||:||||:||||:||||:||||:
Db 571 LICYKESVDQKGQNMDSKRNVLFSVFE 600
```

Wed Oct 22 10:17:50 2003

fviii_arg581k.rapb

Page 6

Search completed: October 21, 2003, 19:01:22
Job time : 21.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:49:52 ; Search time 17.5 Seconds
(without alignments)
164.861 Million cell updates/sec

Title: FVIII_ARG581K

Perfect score: 153

Sequence: 1 LICYKESVDQKGNQIMSDKRNVLPSVFDE 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	98.0	2351	1 E2HU	coagulation factor
2	147	96.1	2133	2 T42763	coagulation factor
3	146	95.4	2319	2 A47004	coagulation factor
4	92	53.6	2183	2 T42764	coagulation factor
5	76	49.7	2211	1 KFB05	coagulation factor
6	71	46.4	2224	1 KFH05	coagulation factor
7	57.5	37.6	1069	1 KUH0	ferroxidase (BC 1
8	55	35.9	742	2 AB2470	photosystem I P700
9	54	35.3	337	2 S73737	glyceradenhyde-3-P
10	50	32.7	150	2 G87590	cytochrome c-type
11	50	32.7	211	2 T34501	hypothetical prote
12	50	32.7	325	2 T42161	J3L protein - vari
13	50	32.7	325	2 S33100	H2R protein - vari
14	50	32.7	325	2 T28524	hypothetical prote
15	49	32.0	324	1 QQVZH3	H3 protein - vacci
16	49	32.0	324	2 C42514	H3L protein - vacc
17	49	32.0	324	2 T37369	IMV membrane assoc
18	49	32.0	641	2 T51385	hypothetical prote
19	48.5	31.7	1078	2 E75407	isoleucyl-tRNA syn
20	48.5	31.7	1817	2 T10689	hypothetical prote
21	48	31.4	301	2 F90115	beta subunit of G
22	48	31.4	467	2 T38353	serine hydroxymeth
23	48	31.4	595	2 F70402	penicillin binding
24	48	31.4	649	2 S48379	glycerol-3-phospha
25	47.5	31.0	472	2 T37918	serine hydroxymeth
26	47.5	31.0	1115	2 A45761	Ca2+-transporting
27	47	30.7	442	2 B81418	adenylosuccinate l
28	47	30.7	976	1 TVHUKT	protein-tyrosine k
29	47	30.7	1571	2 S50669	hypothetical prote

ALIGNMENTS

RESULT 1

E2HU

coagulation factor VIII precursor [validated] - human

N;Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant c

C;Species: Homo sapiens (man)

C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000

C;Accession: I54318; A00525; I58059; A23584; A36174; A42348; A43986; S63527; S66445;

R;Gitschier, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

A;Title: Sequence of the exon-containing regions of the human factor VIII gene.

A;Reference number: I54318; MUID:93265012; PMID:1303178

A;Accession: I54318

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1921,'S','S',1923-2351 <RES>

A;Cross-references: GB:M88648; NID:gl82381; PIDN:AAA52420.1; PID:gl82393

R;Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; See

Nature 312, 330-337, 1984

A;Title: Expression of active human factor VIII from recombinant DNA clones.

A;Reference number: A00525; MUID:85061548; PMID:6438526

A;Accession: A00525

A;Molecule type: mRNA

A;Residues: 1-2351 <R00>

A;Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179

R;Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.

S, D.N.; Hewick, R.M.

Nature 312, 342-347, 1984

A;Title: Molecular cloning of a cDNA encoding human antihemophilic factor.

A;Reference number: I58059; MUID:85061550; PMID:6438528

A;Accession: I58059

A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-74,'V','V',76-1259,'E',1261-2351 <RE2>

A;Cross-references: GB:K01740; NID:gl82802; PIDN:AAA52484.1; PID:gl82803

R;Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.;

B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favaloro, J.; Hansen, J.

DNA 4, 333-349, 1985

A;Title: Characterization of the polypeptide composition of human factor VIII:C and

A;Reference number: A23584; MUID:86081164; PMID:3935400

A;Accession: A23584

A;Molecule type: mRNA

A;Residues: 1-2351 <TRU>

A;Cross-references: GB:M14113; NID:gl82817; PIDN:AAA52485.1; PID:gl82818

R;Eaton, D.; Rodriguez, H.; Vehar, G.A.

Biochemistry 25, 505-512, 1986

A;Title: Proteolytic processing of human factor VIII. Correlation of specific cleava

ity.

A;Reference number: A26174; MUID:86159740; PMID:3082357

A;Accession: A26174

A;Molecule type: protein

A;Residues: 20-36;392-399,'X',401-402;1668-1678;1709-1722,'D',1723-1725;1741-1755 <E

R;Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
A:Title: Identification and functional importance of tyrosine sulfate residues within re
A:Reference number: A42348; PMID:92207952; PMID:1554716
A:Accession: A42348
A:Molecule type: protein
A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
A:Experimental source: recombinant material from Chinese hamster ovary cells
A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
J:Pay, P.J.; Smudzin, T.M.
J. Biol. Chem. 264, 14005-14010, 1989
A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
A:Reference number: A43986; PMID:89340500; PMID:2503509
A:Accession: A43986
A:Molecule type: protein
A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <RAY>
R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
A:Reference number: A56109; PMID:91093266; PMID:1898735
A:Contents: annotation; sulfation
R:Gitschler, J.; Wood, W.L.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A:Title: Characterization of the human factor VIII gene.
A:Reference number: A56196; PMID:85061547; PMID:6438525
A:Contents: annotation; introns
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o
A:Reference number: A56216; PMID:95338127; PMID:7613471
A:Contents: annotation; disulfide bonds
A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
R:Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
A:Reference number: S63527; PMID:96163459; PMID:8575434
A:Accession: S63527
A:Molecule type: protein
A:Residues: 733-752;753-759 <KJ>
R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
Eur. J. Biochem. 232, 19-27, 1995
A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
A:Reference number: S66445; PMID:96048024; PMID:7556150
A:Accession: S66445
A:Molecule type: protein
A:Status: preliminary
A:Residues: 1668-1685 <LN>
A:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
C:Genetics:
A:Gene: GDB:F8C
A:Cross-references: GDB:119124; OMIM:306700
A:Map position: Xq28-Xq28
A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
A:Pathway: blood coagulation
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
F:20-740/Product: coagulation factor VIIIa heavy chain #status experimental <ACH>
F:20-356/Domain: A1 <DA1>
F:23-348/Domain: A2 <DA2>
F:392-759/Domain: A2 <DA2>
F:402-730/Domain: ferroxidase repeat homology <FO1>
F:760-1667/Domain: B <DBO>
F:1668-2351/Product: coagulation factor VIIIa light chain #status experimental <ACL>
F:1709-2038/Domain: A3 <DA3>
F:1716-2038/Domain: ferroxidase repeat homology <FO3>
F:2039-2191/Domain: C1 <DC1>
F:2039-2188/Domain: discoidin I amino-terminal homology <DN1>
F:2192-2351/Domain: C2 <DC2>
F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>
F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #sta
F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
F:365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experiment
F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
F:414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F:2193-2345/Disulfide bonds: #status predicted

Query Match 98.0%; Score 150; DB 1; Length 2351;
Best Local Similarity 96.7%; Pred. No. 5e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

RESULT 2

T42763
coagulation factor VIII precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C:Accession: T42763
R:Lollar, P.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z22269
A:Accession: T42763
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2133 <LOL>
A:Cross-references: EMBL:U49517; NID:gl511633; PTD:gl511634; PIDN:AA06705.1
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 96.1%; Score 147; DB 2; Length 2133;
Best Local Similarity 93.3%; Pred. No. 1.2e-12;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
|||||:|||||:|||||:|||||:|||||
Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600

RESULT 3

A47004
coagulation factor VIII precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
C:Accession: A47004
R:Elder, B.; Laskich, D.; Gitschler, J.
Genomics 16, 374-379, 1993
A:Title: Sequence of the murine factor VIII cDNA.
A:Reference number: A47004; MUID:93300511; PMID:8314577
A:Accession: A47004
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2319 <EID>
A:Cross-references: GB:L05573; NID:q192456; PIDN:AAA37385.1; PID:q192457
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
F:1-19/Domain: signal sequence #status predicted <SIG>
F:23-349/Domain: ferroxidase repeat homology <FO1>
F:402-730/Domain: ferroxidase repeat homology <FO2>
F:1686-2006/Domain: ferroxidase repeat homology <FO3>
F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 95.4%; Score 146; DB 2; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 1.8e-12;
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
 |||||:|||||:|||||:|||||:|||||:|||||
 Db 571 LICYKESVDQKGNQIMSDKRNVLFSVFDE 600

RESULT 4

T42764
 coagulation factor V - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42764
 R:Yang, T.L.; Cui, J.; Rehmulla, A.; Yang, A.; Mousalli, M.; Kaufman, R.J.; Ginsburg, Blood 91, 4593-4599, 1998
 A:Title: The structure and function of murine factor V and its inactivation by protein C
 A:Reference number: Z22270; MUID:98282202; PMID:9616155
 A:Accession: T42764
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2183 <YAN>
 A:Cross-references: EMBL:U52925; NID:G3219690; PID:G3219691; PIDN:AAC99553.1
 C:Function:
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
 C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
 F:350-682/Domain: ferroxidase repeat homology <FOX1>
 F:1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 53.6%; Score 82; DB 2; Length 2183;
 Best Local Similarity 53.3%; Pred. No. 0.003;
 Matches 16; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
 ||||:|||||:|||||:|||||:|||||:|||||
 Db 190 LICKKGTLTEDGQKMFQHVLMFAVFDE 219

RESULT 5

KFB05
 coagulation factor V precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 04-Mar-1993 #sequence_revision 28-Apr-1995 #text_change 11-Jun-1999
 C:Accession: A42580; A36497
 R:Guinto, E.R.; Esmon, C.T.; Mann, K.G.; MacGillivray, R.T.
 J. Biol. Chem. 267, 2971-2978, 1992
 A:Title: The complete cDNA sequence of bovine coagulation factor V.
 A:Reference number: A42580; MUID:92147638; PMID:1737753
 A:Accession: A42580
 A:Molecule type: mRNA
 A:Residues: 1-2211 <GUI>
 A:Cross-references: GB:M81440; NID:G163037; PIDN:AAA30512.1; PID:G163038
 A:Note: sequence extracted from NCBI backbone (NCBIN:80774, NCBI:80776)
 R:Kalafatis, M.; Jenny, R.J.; Mann, K.G.
 J. Biol. Chem. 265, 21580-21589, 1990
 A:Title: Identification and characterization of a phospholipid-binding site of bovine fa
 A:Reference number: A36497; MUID:91072354; PMID:2254316
 A:Accession: A36497
 A:Molecule type: protein
 A:Residues: 1566-1570, 'X', 1572-1581, 'X', 1583-1584; 1673-1676, 'X', 1678-1679, 'X', 1681, 'X', 1
 R:Xue, J.; Kalafatis, M.; Silveira, J.R.; Kung, C.; Mann, K.G.
 Biochemistry 33, 13109-13116, 1994
 A:Title: Determination of the disulfide bridges in factor Va heavy chain.
 A:Reference number: A55979; MUID:95034740; PMID:7947716
 A:Contents: annotation
 A:Note: 566-Cys and 617-Cys were shown to have free sulfhydryls
 C:Comment: factor V is activated by thrombin and partially by coagulation factor Xa.
 C:Function:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor xa prot
 A:Pathway: blood coagulation

C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidas
 C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plas
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-2211/Product: coagulation factor V #status predicted <MAT>
 F:29-741/Product: coagulation factor Va heavy chain #status predicted <VAH>
 F:29-345/Domain: A1 <DAL>
 F:33-329/Domain: ferroxidase repeat homology <FO1>
 F:346-695/Domain: A2 <DA2>
 F:351-688/Domain: ferroxidase repeat homology <FO2>
 F:696-1564/Domain: B <DOB>
 F:1175-1437/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
 F:1565-2211/Product: coagulation factor Va light chain #status predicted <VAL>
 F:1565-1892/Domain: A3 <DA3>
 F:1572-1892/Domain: ferroxidase repeat homology <FO3>
 F:1654-1752/Region: phospholipid binding #status predicted
 F:1893-2051/Domain: C1 <DC1>
 F:1893-2048/Domain: discoidin I amino-terminal homology <DNI>
 F:2052-2211/Domain: C2 <DC2>
 F:2052-2208/Domain: discoidin I amino-terminal homology <DN2>
 F:167-193,248-329,499-525/Disulfide bonds: #status experimental
 F:225-239,297,382,460,553,587,774,780,902,952,964,1044,1053,1062,1071,1078,1
 F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted
 F:363-697/1537/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
 F:533-534/Cleavage site: Arg-Gly (protein C) #status predicted
 F:607-688,1712-1738,1894-2048,2053-2208/Disulfide bonds: #status predicted
 F:741-742/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
 F:1034-1035/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predict
 F:1564-1565/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 49.7%; Score 76; DB 1; Length 2211;
 Best Local Similarity 46.7%; Pred. No. 0.022;
 Matches 14; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30
 ||||:|||||:|||||:|||||:|||||:|||||
 Db 191 LICKKGTLTEDGQKMFQHVLMFAVFDE 220

RESULT 6

KFB05
 coagulation factor V precursor [validated] - human
 N:Alternate names: coagulation labile factor; proaccelerin
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 08-Dec-2000
 C:Accession: A56172; A42344; A28028; A27498; A25897
 R:Cripe, L.D.; Moore, K.D.; Kane, W.H.
 Biochemistry 31, 3777-3785, 1992
 A:Title: Structure of the gene for human coagulation factor V.
 A:Reference number: A42344; MUID:92232668; PMID:1567832
 A:Accession: A56172
 A:Molecule type: DNA
 A:Residues: 1-2224 <CRI>
 A:Cross-references: GB:J05368
 A:Accession: A42344
 A:Molecule type: DNA
 A:Residues: 48-58;79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-
 2070;2111-2120;2172-2181 <CR2>
 R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; K
 Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
 A:Title: Complete cDNA and derived amino acid sequence of human factor V.
 A:Reference number: A28028; MUID:87260886; PMID:3110773
 A:Accession: A28028
 A:Molecule type: mRNA
 A:Residues: 1-857, 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <
 A:Cross-references: GB:M16967
 A:Note: parts of this sequence, including the amino end of the mature protein, were
 R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
 Biochemistry 26, 6508-6514, 1987
 A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of
 A:Reference number: A27498; MUID:88107560; PMID:2827731
 A:Accession: A27498
 A:Molecule type: mRNA

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <SHC>
 A:Cross-references: GB:X16780; NID:g5830555; PIDN:CAB54686.1; PID:g5830647
 A:Experimental source: strain Garcia-1966
 C:Genetics:
 A:Gene: i3L
 C:Superfamily: vaccinia virus H3 protein

Query Match 32.7%; Score 50; DB 2; Length 325;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 DQKGNQIMSDKRNVL 24
 |||:::||||:
 Db 37 DVKDNEVMPKRNVI 52

RESULT 13

S33100
 H2R protein - variola virus
 N:Alternate names: I3L protein
 C:Species: variola virus
 C>Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 23-Mar-2001
 C:Accession: S33100; B36846
 R:Shchelkunov, S.N.; Blinov, V.M.; Totmenin, A.V.; Marennikova, S.S.; Kolykhalov, A.A.;
 dzhaparidze, O.G.; Sandakhchiev, L.S.
 A:Title: Nucleotide sequence analysis of variola virus HindIII M, L, I genome fragments.
 A:Reference number: S33069; MUID:93190624; PMID:8383392
 A:Accession: S33100
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-325 <SHC>
 A:Cross-references: EMBL:X67119; NID:g62330; PIDN:CAA47585.1; PID:g62362
 A:Experimental source: strain India-1967, isolate Ind3
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992
 R:Blinov, V.M.
 submitted to GenBank, November 1992

A:Reference number: A36859
 A:Accession: B36846
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <BLI>
 A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49027.1; PID:g297266
 A:Experimental source: strain India-1967, ssp. major, isolate Ind3
 C:Superfamily: vaccinia virus H3 protein

Query Match 32.7%; Score 50; DB 2; Length 325;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 DQKGNQIMSDKRNVL 24
 |||:::||||:
 Db 37 DVKDNEVMPKRNVI 52

RESULT 14

T28524
 hypothetical protein I3L - variola major virus
 C:Species: variola major virus
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T28524
 R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
 Nature 366, 748-751, 1993
 A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
 A:Reference number: Z20488; MUID:94068747; PMID:8264798
 A:Accession: T28524
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-325 <MAS>
 A:Cross-references: EMBL:L22579; NID:g623595; PIDN:AAA60834.1; PID:g439004
 A:Experimental source: strain Bangladesh-1975

C:Superfamily: vaccinia virus H3 protein

Query Match 32.7%; Score 50; DB 2; Length 325;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 DQKGNQIMSDKRNVL 24
 |||:::||||:
 Db 37 DVKDNEVMPKRNVI 52

RESULT 15

QQVZB3
 H3 protein - vaccinia virus (strain WR)
 C:Species: vaccinia virus
 C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
 C:Accession: C24481
 R:Roedel, J.L.; Earl, P.L.; Weir, J.P.; Moss, B.
 J. Virol. 60, 436-449, 1986
 A:Title: Conserved TAAATG sequence at the transcriptional and translational initiation
 t.
 A:Reference number: A93022; MUID:87036903; PMID:3021979
 A:Accession: C24481
 A:Molecule type: DNA
 A:Residues: 1-324 <ROS>
 A:Cross-references: GB:M13209; NID:g335739; PIDN:AAB59838.1; PID:g335743
 C:Superfamily: vaccinia virus H3 protein
 C:Keywords: late protein

Query Match 32.0%; Score 49; DB 1; Length 324;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 DQKGNQIMSDKRNVL 24
 |||:::||||:
 Db 36 DVKDNEVMPKRNVI 51

Search completed: October 21, 2003, 18:58:44
 Job time : 20.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:22 ; Search time 9.25 Seconds
(without alignments)
152.519 Million cell updates/sec

Title: FVIII_ARG581K

Perfect score: 153

Sequence: 1 LICIKESVDQGNQIMSKRNVILFSVFDE 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	150	98.0	2351	1	FAB_HUMAN	P00451 homo sapien
2	147	96.1	2133	1	FAB_PIG	P12263 sus scrofa
3	146	95.4	2219	1	FAB_MOUSE	Q06194 mus musculus
4	76	49.7	2211	1	FAS_BOVIN	Q28107 bos taurus
5	74	48.4	2258	1	FAS_PIG	Q9gip1 sus scrofa
6	71	46.4	2224	1	FAS_HUMAN	P12259 homo sapien
7	57.5	37.6	1065	1	CERU_HUMAN	P00450 homo sapien
8	55	35.9	741	1	PAB2_HUMAN	Q8yli4 anabaena sp
9	54	35.3	337	1	G3P_MYCPN	P75358 mycoplasma
10	53.5	35.0	185	1	RRE_WIGBR	Q8d2g5 wiggleswort
11	50	32.7	325	1	VP35_VAVR	P33059 variola vir
12	49	32.0	324	1	VP35_VACCC	P20497 vaccinia vi
13	49	32.0	324	1	VP35_VACCC	P07240 vaccinia vi
14	48	31.4	467	1	GLYD_SCHPO	O13972 schizosacch
15	48	31.4	649	1	GPDM_YEAST	P32191 saccharomyc
16	47.5	31.0	452	1	TBC_PLAFO	P34787 plasmodium
17	47.5	31.0	472	1	GLXC_SCHPO	Q10104 schizosacch
18	47	30.7	734	1	PSAB_OENHO	Q9mtn7 oenothera h
19	47	30.7	976	1	KIT_HUMAN	P10721 homo sapien
20	47	30.7	1571	1	ATC5_YEAST	P32660 saccharomyc
21	46.5	30.4	214	1	VAID_DESSY	O06506 desulfuroco
22	46.5	30.4	451	1	GLYA_ARCFU	O29406 archaeoglob
23	46	30.1	322	1	PL16_ARATH	P48486 arabidopsis
24	46	30.1	440	1	CLPX_AZOVI	P33683 azotobacter
25	46	30.1	644	1	YNL5_YEAST	P53925 saccharomyc
26	46	30.1	660	1	CIR4_BOVIN	Q05037 bos taurus
27	46	30.1	661	1	PSAB_PROHO	P58387 prochloroth
28	46	30.1	730	1	PSAB_STNY3	P29255 synechocyst
29	46	30.1	732	1	PSAB_STNY2	P17155 synechococ
30	46	30.1	733	1	PSAB_ODOSI	P49480 odontella s
31	46	30.1	734	1	PSAB_ANTMA	Q33332 antirrhinum
32	46	30.1	734	1	PSAB_ARATH	P56767 arabidopsis
33	46	30.1	734	1	PSAB_ATRBE	Q8s8x5 atropa bell

34	46	30.1	734	1	PSAB_CHLYU	P56342 chlorella v
35	46	30.1	734	1	PSAB_CYACA	Q9tiq6 cyanidium c
36	46	30.1	734	1	PSAB_EUGGR	P19431 euglena gra
37	46	30.1	734	1	PSAB_GUITH	O78507 guillardia
38	46	30.1	734	1	PSAB_LOTJA	P58385 lotus japon
39	46	30.1	734	1	PSAB_MESVI	Q9mur7 mesostigma
40	46	30.1	734	1	PSAB_NEPOL	Q9tkw1 nephiroselmi
41	46	30.1	734	1	PSAB_ORYSA	P12156 oryza sativ
42	46	30.1	734	1	PSAB_PEA	P05311 pisum sativ
43	46	30.1	734	1	PSAB_PINTH	P41640 pinus thunb
44	46	30.1	734	1	PSAB_PORPU	P51285 porphyra pu
45	46	30.1	734	1	PSAB_SPIOL	P06512 spinachia ol

ALIGNMENTS

RESULT 1

ID	FAB_HUMAN	STANDARD;	PRT;	2351 AA.
AC	P00451;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Coagulation factor VIII precursor (Procoagulant component)			
DE	(Antihemophilic factor) (AHF).			
GN	P8 OR F8C.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86081164; PubMed=3935400;			
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,			
RA	Hartog K., Kuo C.H., Maslarz F.R., Merryweather J.P., Najarian R.,			
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,			
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,			
RA	Nordfang O., Ezban M.,			
RT	"Characterization of the polypeptide composition of human factor VIII;C and the nucleotide sequence and expression of the human kidney cDNA.";			
RT	DNA 4:333-349(1985).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85061548; PubMed=64385326;			
RA	Wood W.L., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,			
RA	Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,			
RA	Delwart E., Tuddenham E.G.D., Vohar G.A., Lawn R.M.;			
RT	"Expression of active human factor VIII from recombinant DNA clones.";			
RT	Nature 312:330-337(1984).			
RL	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85061550; PubMed=64385328;			
RA	Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,			
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,			
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,			
RA	Hewick R.M.;			
RT	"Molecular cloning of a cDNA encoding human antihemophilic factor.";			
RT	Nature 312:342-347(1984).			
RL	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93265012; PubMed=1303178;			
RA	Gitschier J., Wood W.L.;			
RT	"Sequence of the exon containing regions of the human factor VIII gene.";			
RT	Hum. Mol. Genet. 1:199-200(1992).			
RL	[5]			
RP	SEQUENCE OF 2064-2070 FROM N.A.			
RA	de Water N.S., Williams R., Browett P.J.;			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RL	[6]			
RP	SULFATION OF TYR-1699.			

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RX SULFATION.
 RP MEDLINE=92207952; PubMed=1554716;
 RA Pittman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RX STRUCTURE BY NMR OF 2322-2343.
 RP MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RP MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RP MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:11-12(1989).
 RN [11]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RP MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RX VARIANT HEMA GLN-2326.
 RP MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac";
 RL Science 232:1415-1416(1986).
 RN [13]
 RX VARIANT HEMA PRO-2135.
 RP MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RX VARIANT HEMA GLN-2228.
 RP MEDLINE=86191889; PubMed=2833855;
 RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RX VARIANT HEMA GLY-291.
 RP MEDLINE=86220354; PubMed=2835904;
 RA Yousoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RX VARIANT HEMA CYS-1708.
 RP MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1.689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RX VARIANT HEMA CYS-391.
 RP MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RX VARIANT HEMA LEU-189.
 RP MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RX VARIANT HEMA LEU-2326.
 RP MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RX VARIANT HEMA HIS-391.
 RP MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RX VARIANT HEMA CYS-1708.
 RP MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RX VARIANTS HEMA GLN-2228 AND LEU-2326.
 RP MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Marzu S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RX VARIANT HEMA CYS-391.
 RP MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RX VARIANTS HEMA PHE-1699 AND CYS-1708.
 RP MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:55-71(1990).
 RN [25]
 RX VARIANTS HEMA CYS-1728 AND ASP-1941.
 RP MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

```

Query Match          98.0%; Score 150; DB 1; Length 2351;
Best Local Similarity 96.7%; Pred. No. 2.9e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSKRNVLFSVFDE 30
    |||||:|||||:|||||:|||||:|||||
DB 571 LICYKESVDQGNQIMSKRNVLFSVFDE 600

RESULT 2
FAS_PIG
ID FAS_PIG STANDARD; PRT; 2133 AA.
AC P12263; Q95243;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
GN F8 OR CF8.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RA Healey J.F., Lubin I.M., Lollar P.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE OF 705-1573 FROM N.A.
RX MEDLINE=86287369; PubMed=3016730;
RA Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA Kaufman R.J.;
RT "A large region (approximately equal to 95 kDa) of human factor VIII
RT is dispensable for in vitro procoagulant activity.";
RN [3]
Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
RX SEQUENCE OF 392-759 FROM N.A.
RA MEDLINE=94179260; PubMed=7510693;
RA Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT "Elimination of a major inhibitor epitope in factor VIII.";
RN [4]
J. Biol. Chem. 269:8639-8641(1994).
CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
CC ACTIVATED FORM, FACTOR XA.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; U49517; AAB06705.1; .
DR PIR; A25945; A25945.
DR PIR; T42763; T42763.
DR HSSP; P00451; 1CRG.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR000421; FA58-C.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
KW Signal; Glycoprotein; Sulfation.
FT SIGNAL 1 19 POTENTIAL.

```

```

FT CHAIN 20 2133 COAGULATION FACTOR VIII.
FT DOMAIN 20 357 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 207 357 PLASTOCYANIN-LIKE 2.
FT DOMAIN 399 730 F5/8 TYPE A 2.
FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
FT DOMAIN 760 1599 B.
FT DOMAIN 1495 1822 F5/8 TYPE A 3.
FT DOMAIN 1495 1659 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1669 1822 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1822 1970 F5/8 TYPE C 1.
FT DOMAIN 1975 2127 F5/8 TYPE C 2.
FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1449 1450 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT SITE 1490 1491 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT MOD_RES 737 737 SULFATION (BY SIMILARITY).
FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
FT DISULFID 173 199 PROBABLE.
FT DISULFID 547 573 PROBABLE.
FT DISULFID 1633 1659 PROBABLE.
FT DISULFID 1822 1970 BY SIMILARITY.
FT DISULFID 1975 2127 BY SIMILARITY.
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 929 929 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 985 985 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1111 1111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1181 1181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1208 1208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1245 1245 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1335 1335 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1611 1611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1919 1919 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 713 713 N -> M (IN REF. 2).
FT CONFLICT 734 734 I -> T (IN REF. 2).
FT CONFLICT 792 792 G -> Q (IN REF. 2).
FT CONFLICT 1133 1133 E -> F (IN REF. 2).
FT CONFLICT 1191 1191 I -> L (IN REF. 2).
FT CONFLICT 1209 1209 R -> F (IN REF. 2).
FT CONFLICT 1437 1437 C -> G (IN REF. 2).
FT CONFLICT 1456 1456 F -> R (IN REF. 2).
FT CONFLICT 1539 1539 F -> R (IN REF. 2).
FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
SQ SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;

Query Match          96.1%; Score 147; DB 1; Length 2133;
Best Local Similarity 93.3%; Pred. No. 7e-13;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQGNQIMSKRNVLFSVFDE 30
    |||||:|||||:|||||:|||||:|||||
DB 571 LICYKESVDQGNQIMSKRNVLFSVFDE 600

RESULT 3
FAS_MOUSE
ID FAS_MOUSE STANDARD; PRT; 2319 AA.
AC Q06194;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
GN F8 OR CF8 OR F8C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.

```

OX NCBI_TaxID=10090;
 RP [1]
 RN SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=93300511; PubMed=8314577;
 RA Elder B., Lakich D., Gitschier J.;
 RT "Sequence of the murine factor VIII cDNA";
 RL Genomics 16:374-379(1993).
 CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR XA.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L05573; AAA37385.1; -
 DR PIR; A47004; A47004.
 DR HSPD; P00451; ICFG.
 DR MGD; MGI:88383; F8.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR000421; FA58_C.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR Pfam; PF00754; F5_F8_type_C; 2.
 DR SMART; SM00231; FA58C; 2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR PROSITE; PS00022; FA58C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 KW Signal; Glycoprotein; Sulfation.
 FT SIGNAL 19 POTENTIAL.
 FT CHAIN 20 2319 COAGULATION FACTOR VIII.
 FT DOMAIN 20 349 F5/8 TYPE A 1.
 FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 207 349 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 399 730 F5/8 TYPE A 2.
 FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 760 1640 B.
 FT DOMAIN 1683 2008 F5/8 TYPE A 3.
 FT DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.
 FT DOMAIN 1835 2008 PLASTOCYANIN-LIKE 6.
 FT DOMAIN 2008 2156 F5/8 TYPE C 1.
 FT DOMAIN 2161 2313 F5/8 TYPE C 2.
 FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT MOD_RES 367 367 SULFATION (BY SIMILARITY).
 FT MOD_RES 737 737 SULFATION (BY SIMILARITY).
 FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
 FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
 FT MOD_RES 1669 1669 SULFATION (REQUIRED FOR VWF BINDING) (BY SIMILARITY).
 FT MOD_RES 1687 1687 SULFATION (BY SIMILARITY).
 FT DISULFID 173 199 PROBABLE.
 FT DISULFID 547 573 PROBABLE.
 FT DISULFID 1819 1345 PROBABLE.
 FT DISULFID 2008 2156 BY SIMILARITY.
 FT DISULFID 2161 2313 BY SIMILARITY.
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1136 1136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1161 1161 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1192 1192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;
 Query Match 95.4%; Score 146; DB 1; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 1.1e-12;
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LICYKESVDQGNQIMSDKRNVLFSVFDE 30
 Db 571 LICYKESVDQGNQIMSDKRNVLFSVFDE 600
 RESULT 4
 FA5_BOVIN STANDARD; PRT; 2211 AA.
 ID FA5_BOVIN Q28108;
 AC Q28107; Q28108;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Coagulation factor V precursor (Activated protein C cofactor).
 GN F5.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92147638; PubMed=1737753;
 RA Quinto E.R., Esmon C.T., Mann K.G., Macgillivray R.T.;
 RT "The complete cDNA sequence of bovine coagulation factor V";
 RL J. Biol. Chem. 267:2971-2978(1992).
 CC -!- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
 CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
 CC -!- SUBUNIT: Factor Va is composed of a heavy chain and a light
 CC chain, noncovalently bound. The interaction between the two chains
 CC is calcium-dependent.
 CC -!- DOMAIN: Domain B contains 29.5 X 9 AA tandem repeats, and 2 X 14
 CC AA repeats.
 CC -!- PTM: Thrombin activates factor V proteolytically to the active
 CC cofactor, factor Va (formation of a heavy chain at the N-
 CC terminus and a light chain at the C-terminus).
 CC -!- PTM: SULFATION IS REQUIRED FOR EFFICIENT THROMBIN CLEAVAGE AND
 CC ACTIVATION AND FOR FULL PROCOAGULANT ACTIVITY (BY SIMILARITY).
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

DR EMBL; M81440; AAA30512.1; -
DR EMBL; M81441; AAA30513.1; -
DR PIR; A42580; KFB05.
DR HSSP; P12259; LCZT.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C.1; 2.
DR PROSITE; PS01286; FA58C.2; 2.
DR PROSITE; PS50022; FA58C.3; 2.
DR PROSITE; PS50027; MULTICOPPER_OXIDASE; 2.
KW Blood coagulation; Glycoprotein; Sulfation; Calcium; Signal; Zymogen;
KW Repeat.
FT SIGNAL 1 28
FT CHAIN 29 2211
FT CHAIN 29 741
FT PEPTIDE 742 1564
FT CHAIN 1565 2211
FT DOMAIN 30 327
FT DOMAIN 30 193
FT DOMAIN 203 327
FT DOMAIN 348 686
FT DOMAIN 348 525
FT DOMAIN 535 686
FT DOMAIN 696 1564
FT DOMAIN 1124 1151
FT REPEAT 1124 1137
FT REPEAT 1138 1151
FT DOMAIN 1188 1453
FT REPEAT 1188 1196
FT REPEAT 1197 1205
FT REPEAT 1206 1214
FT REPEAT 1215 1223
FT REPEAT 1224 1232
FT REPEAT 1233 1241
FT REPEAT 1242 1250
FT REPEAT 1251 1259
FT REPEAT 1260 1268
FT REPEAT 1269 1277
FT REPEAT 1278 1286
FT REPEAT 1287 1295
FT REPEAT 1296 1304
FT REPEAT 1305 1313
FT REPEAT 1314 1322
FT REPEAT 1323 1331
FT REPEAT 1332 1340
FT REPEAT 1341 1349
FT REPEAT 1350 1358
FT REPEAT 1359 1367
FT REPEAT 1368 1376
FT REPEAT 1377 1385
FT REPEAT 1386 1394
FT REPEAT 1395 1403
FT REPEAT 1404 1412
FT REPEAT 1413 1421
FT REPEAT 1422 1430
FT REPEAT 1431 1439
FT REPEAT 1440 1444
FT REPEAT 1445 1453
FT DOMAIN 1569 1890
FT DOMAIN 1569 1738
FT DOMAIN 1748 1890
  
```

```

FT DOMAIN 1894 2048
FT DOMAIN 2053 2208
FT SITE 741 742
FT SITE 1034 1035
FT SITE 1564 1565
FT DISULFID 167 193
FT DISULFID 499 525
FT DISULFID 1712 1738
FT DISULFID 1894 2048
FT DISULFID 2053 2208
FT MOD_RES 697 697
FT MOD_RES 701 701
FT MOD_RES 730 730
FT MOD_RES 1513 1513
FT MOD_RES 1529 1529
FT MOD_RES 1537 1537
FT MOD_RES 1541 1541
FT CARBOHYD 225 225
FT CARBOHYD 239 239
FT CARBOHYD 297 297
FT CARBOHYD 382 382
FT CARBOHYD 460 460
FT CARBOHYD 553 553
FT CARBOHYD 587 587
FT CARBOHYD 745 745
FT CARBOHYD 756 756
FT CARBOHYD 774 774
FT CARBOHYD 780 780
FT CARBOHYD 902 902
FT CARBOHYD 952 952
FT CARBOHYD 964 964
FT CARBOHYD 1044 1044
FT CARBOHYD 1053 1053
FT CARBOHYD 1062 1062
FT CARBOHYD 1071 1071
FT CARBOHYD 1078 1078
FT CARBOHYD 1094 1094
FT CARBOHYD 1451 1451
FT CARBOHYD 1490 1490
FT CARBOHYD 1550 1550
FT CARBOHYD 1690 1690
FT CARBOHYD 1839 1839
FT CARBOHYD 1997 1997
FT CARBOHYD 2196 2196
FT VARIANT 587 592
SQ SEQUENCE 2211 AA; 248981 MW; CBBF90H738667C45 CRC64;

Query Match 49.7%; Score 76; DB 1; Length 2211;
Best Local Similarity 46.7%; Pred. No. 0.011;
Matches 14; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 LICKESVDQKGNQIMSDKRNILSVFDE 30
   ||| |::: | | |::: ||||
Db 191 LICKKGLTLEDGTQKMFQKQVLMFAVEDE 220

RESULT 5
FA5_PIG
ID FA5_PIG STANDARD; PRT; 2258 AA.
AC Q9GLPI;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN F5.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING OF F5/8 TYPE A AND C
RC DOMAINS.
RC TISSUE=Liver;
  
```


FT CARBOHYD 1480 1480 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1537 1537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1597 1597 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1737 1737 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1886 1886 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2044 2044 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2243 2243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 2258 AA; 256078 MW; 9159B9B0076A2ACC CRC64;

Query Match 48.4%; Score 74; DB 1; Length 2258;
 Best Local Similarity 50.08; Pred. No. 0.022;
 Matches 15; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LICKVESVDKQGMQIMSDKRNVLFSVFDE 30
 ||| | : : | | | : : : |||
 Db 191 LICKKGLTDEGIQKMFQVLMFAVFDE 220

RESULT 6

FA5_HUMAN

ID FA5_HUMAN STANDARD; PRT; 2224 AA.
 AC P12259; Q14285;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Coagulation factor V precursor (Activated protein C cofactor).
 GN F5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT MET-1764.
 RX MEDLINE=92232668; PubMed=1567832;
 RA Cripe L.D., Moore K.D., Kane W.H.;
 RT "Structure of the gene for human coagulation factor V.";
 RL Biochemistry 31:3777-3785(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT GLU-925.
 RX MEDLINE=87260886; PubMed=3110773;
 RA Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A.,
 RA Hewick R.M., Kaufman R.J., Mann K.G.;
 RT "Complete cDNA and derived amino acid sequence of human factor V.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
 RN [3]
 RP SEQUENCE OF 1-1600 FROM N.A., AND VARIANTS GLU-925 AND ILE-1285.
 RX MEDLINE=88107560; PubMed=2827731;
 RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
 RT "Cloning of cDNAs coding for the heavy chain region and connecting
 RT region of human factor V, a blood coagulation factor with four types
 RT of internal repeats.";
 RL Biochemistry 26:6508-6514(1987).
 RN [4]
 RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
 RX MEDLINE=86313665; PubMed=3092220;
 RA Kane W.H., Davie E.W.;
 RT "Cloning of a cDNA coding for human factor V, a blood coagulation
 RT factor homologous to factor VIII and ceruloplasmin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A.
 RP TISSUE=Fibroblast;
 RX MEDLINE=93203619; PubMed=8454869;
 RA Shen N.L., Fan S.-T., Pyati J., Graff R., Lapolla R.J.,
 RA Edgington T.S.;
 RT "The serine protease cofactor factor V is synthesized by
 RT lymphocytes.";
 RL J. Immunol. 150:2992-3001(1993).
 RN [6]
 RP SULFATION.
 RX MEDLINE=94264012; PubMed=8204629;
 RA Pittman D.D., Tomkinson K.N., Michnick D., Selighsohn U.,
 RA Kaufman R.J.;

RT "Posttranslational sulfation of factor V is required for efficient
 RT thrombin cleavage and activation and for full procoagulant activity.";
 RL Biochemistry 33:6952-6959(1994).
 RN [7]
 RP SULFATION.
 RX MEDLINE=90366699; PubMed=2168225;
 RA Hortic G.L.;
 RT "Sulfation of tyrosine residues in coagulation factor V.";
 RL Blood 76:946-952(1990).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2065-2224.
 RX MEDLINE=20052169; PubMed=10586886;
 RA Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M.A., Kim S.W.,
 RA Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
 RA Fuentes-Prior P.;
 RT "Crystal structures of the membrane-binding C2 domain of human
 RT coagulation factor V.";
 RL Nature 402:434-439(1999).
 RN [9]
 RP VARIANT MET-1764.
 RX MEDLINE=95179146; PubMed=7874144;
 RA Bayston T.A., Ireland H., Olds R.J., Thein S.L., Lane D.A.;
 RT "A polymorphism in the human coagulation factor V gene.";
 RL Hum. Mol. Genet. 3:2085-2085(1994).
 RN [10]
 RP VARIANT APCR GLN-534.
 RX MEDLINE=94217810; PubMed=8164741;
 RA Bertina R.M., Koelman B.P.C., Koster T., Rosendaal F.R.,
 RA Dirven R.J., de Ronde H., van der Velden P.A., Reitsma P.H.;
 RT "Mutation in blood coagulation factor V associated with resistance to
 RT activated protein C.";
 RL Nature 369:64-67(1994).
 RN [11]
 RP VARIANTS ILE-1285 AND ARG-1327.
 RX MEDLINE=96351768; PubMed=8713778;
 RA Lunghi B., Iacoviello L., Gemmati D., Dilasio M.G., Castoldi E.,
 RA Pinotti M., Castaman G., Redaelli R., Mariani G., Marchetti G.,
 RA Bernardi F.;
 RT "Detection of new polymorphic markers in the factor V gene:
 RT association with factor V levels in plasma.";
 RL Thromb. Haemost. 75:45-48(1996).
 RN [12]
 RP VARIANT APCR GLY-334, AND VARIANT IYS-513.
 RX MEDLINE=98122763; PubMed=9454741;
 RA Chan W.P., Lee C.K., Kwong Y.L., Lam C.K., Liang R.;
 RT "A novel mutation of Arg306 of factor V gene in Hong Kong Chinese.";
 RL Blood 91:1135-1139(1998).
 RN [13]
 RP VARIANT APCR THR-334.
 RX MEDLINE=98122764; PubMed=9454742;
 RA Williamson D., Brown K., Luddington R., Baglin C., Baglin T.;
 RT "Factor V Cambridge: a new mutation (Arg306-to-Thr) associated with
 RT resistance to activated protein C.";
 RL Blood 91:1140-1144(1998).
 RN [14]
 RP VARIANTS HIS-107; THR-413; IYS-513; SER-809; THR-817; ARG-858;
 RX ARG-865; GLU-925; GLN-1146; ALA-1530; SER-1685; VAL-1749; MET-1764;
 RP ILE-1820 AND GLY-2222, AND VARIANT APCR GLN-534.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [15]
 RP ERRATUM.
 RX Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).

Proc. Natl. Acad. Sci. U.S.A. 78:790-794(1981).
 [8]
 RP SEQUENCE OF 907-1065.
 RA MEDLINE=80137543; PubMed=6987229;
 RX Kingston I.B., Kingston B.L., Putnam F.W.;
 RT "Primary structure of a histidine-rich proteolytic fragment of human
 ceruloplasmin. I. Amino acid sequence of the cyanogen bromide
 peptides.";
 RL J. Biol. Chem. 255:2878-2885(1980).
 RN [9]
 RP SEQUENCE OF 907-1065.
 RX MEDLINE=80137544; PubMed=6987230;
 RA Kingston I.B., Kingston B.L., Putnam F.W.;
 RT "Primary structure of a histidine-rich proteolytic fragment of human
 ceruloplasmin. II. Amino acid sequence of the tryptic peptides.";
 RL J. Biol. Chem. 255:2886-2896(1980).
 RN [10]
 RP SEQUENCE OF 1007-1061 FROM N.A.
 RX MEDLINE=90285218; PubMed=2355023;
 RA Yang F.M., Friedrichs W.E., Cupples R.L., Banifacio M.J.,
 RA Sanford J.A., Horton W.A., Bowman B.H.;
 RT "Human ceruloplasmin. Tissue-specific expression of transcripts
 produced by alternative splicing.";
 RL J. Biol. Chem. 265:10780-10785(1990).
 RN [11]
 RP REVIEW.
 RX MEDLINE=22049919; PubMed=12055353;
 RA Hellman N.E., Gitlin J.D.;
 RT "Ceruloplasmin metabolism and function.";
 RL Annu. Rev. Nutr. 22:439-458(2002).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
 RA Zaitseva I., Zaitsev V., Card G., Moshkov K., Bax B., Ralph A.,
 RA Lindley P.;
 RT "The X-ray structure of human serum ceruloplasmin at 3.1 A: nature of
 the copper centres.";
 RL J. Biol. Inorg. Chem. 1:15-23(1996).
 CC -1- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
 MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
 ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
 AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: 4 Fe(2+) + 4 H(+) + O(2) -> 4 Fe(3+) + 2 H(2)O.
 CC -1- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
 THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
 KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
 BINUCLEAR.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE
 PLASMA.
 CC -1- DISEASE: Defects in CP are the cause of aceruloplasminemia
 [MIM:604290], an autosomal recessive disorder of iron metabolism.
 CC It is characterized by iron accumulation in the brain as well as
 CC visceral organs. Clinical features consist of the triad of retinal
 CC degeneration, diabetes mellitus and neurological disturbances.
 CC -1- DISEASE: Ceruloplasmin levels are decreased in Wilson's disease,
 CC in which copper cannot be incorporated into ceruloplasmin in
 CC liver because of defects in the copper-transporting ATPase 2.
 CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M13699; AAA51976.1; -;
 DR EMBL; D45045; BAA08084.1; -;
 DR EMBL; D45044; BAA08084.1; -;
 DR EMBL; D45028; BAA08084.1; JOINED.
 DR EMBL; D45029; BAA08084.1; JOINED.
 DR EMBL; D45030; BAA08084.1; JOINED.
 DR EMBL; D45031; BAA08084.1; JOINED.

DR EMBL; D45032; BAA08084.1; JOINED.
 DR EMBL; D45033; BAA08084.1; JOINED.
 DR EMBL; D45034; BAA08084.1; JOINED.
 DR EMBL; D45035; BAA08084.1; JOINED.
 DR EMBL; D45036; BAA08084.1; JOINED.
 DR EMBL; D45037; BAA08084.1; JOINED.
 DR EMBL; D45038; BAA08084.1; JOINED.
 DR EMBL; D45039; BAA08084.1; JOINED.
 DR EMBL; D45040; BAA08084.1; JOINED.
 DR EMBL; D45041; BAA08084.1; JOINED.
 DR EMBL; D45042; BAA08084.1; JOINED.
 DR EMBL; D45043; BAA08084.1; JOINED.
 DR EMBL; D00025; BAA00019.1; -;
 DR EMBL; X04135; CAA27752.1; -;
 DR EMBL; X04136; CAA27753.1; -;
 DR EMBL; X04137; CAA27754.1; -;
 DR EMBL; X04138; CAA27755.1; -;
 DR EMBL; M13536; AAA51975.1; -;
 DR EMBL; J05506; -; NOT_ANNOTATED_CDS.
 DR PIR; A25443; KUH0.
 DR PDB; 1K0W; 12-FEB-97.
 DR GlycoSuiteDB; P00450; -;
 DR SWISS-2DPAGE; P00450; HUMAN.
 DR Siena-2DPAGE; P00450; -;
 DR GeneW; HGNC:2295; CP.
 DR MIM; 117700; -;
 DR GO; 604290; -;
 DR GO; 0005615; C:extracellular space; TAS.
 DR GO; 0006879; P:iron ion homeostasis; TAS.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR002355; Multicu_oxidase2.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 DR Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
 KW Signal: Polymorphism: 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1065 CERULOPLASMIN.
 FT DOMAIN 20 357 F5/8 TYPE A 1.
 FT DOMAIN 20 200 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 209 357 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 370 718 F5/8 TYPE A 2.
 FT DOMAIN 370 560 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 570 718 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 730 1061 F5/8 TYPE A 3.
 FT DOMAIN 730 900 PLASTOCYANIN-LIKE 5.
 FT DOMAIN 908 1061 PLASTOCYANIN-LIKE 6.
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 762 762 N-LINKED (GLCNAC. . .).
 FT DISULFID 174 200 PROBABLE.
 FT DISULFID 276 357 PROBABLE.
 FT DISULFID 534 560 PROBABLE.
 FT DISULFID 637 718 PROBABLE.
 FT DISULFID 874 900 PROBABLE.
 FT METAL 120 120 COPPER (TYPE 2) (BY SIMILARITY).
 FT METAL 122 122 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 180 180 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 182 182 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 994 994 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 997 997 COPPER (TYPE 2) (BY SIMILARITY).
 FT METAL 999 999 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 1039 1039 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 1040 1040 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 1041 1041 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 1045 1045 COPPER (TYPE 1) (BY SIMILARITY).

Query Match 37.6%; Score 57.5; DB 1; Length 1065;
 Best Local Similarity 40.9%; Pred. No. 2.4;
 Matches 12; Conservative 10; Mismatches 7; Indels 1; Gaps 1;

QY 1 LICYKESVDQKGNQIMSKRNVLFSVDE 30

```

Db 198 IICKKDSLDKEKHI-DREFFVWFVSVD 226
      ||| |||::: : : : : ||| |||
      198 IICKKDSLDKEKHI-DREFFVWFVSVD 226

RESULT 8
PAB2_ANASP STANDARD; PRT; 741 AA.
ID Q8YL14;
AC Q8YL14;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE photosystem I P700 chlorophyll A apoprotein A2 2 (PsaB 2).
GN PSAB2 OR ALR5314.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RA "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: PsaA and PsaB bind P700, the primary electron donor of
CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
CC oxidoreductase.
CC -!- COFACTOR: PSI electron transfer chain: Binds 6 chlorophyll A, 2
CC phytylquinones and 3 4Fe-4S clusters. PSI core antenna: Binds 90
CC chlorophyll A, 22 carotenoids, 3 phospholipids and 1 galactolipid
CC (By similarity). P700 is a chlorophyll A dimer, A0 is chlorophyll
CC A. A1 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
CC -!- SUBUNIT: A PsaA/B heterodimer binds the P700 chlorophyll special
CC pair and subsequent electron acceptors. The cyanobacterial PSI
CC reaction center (RC) is composed of one copy each of
CC PsaA,B,C,D,E,F,I,J,K,L,M and X, and forms trimeric complexes.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Cellular
CC thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP003599; BAB77013.1; -
CC PIR: AB2470; AB2470.
CC HAMAP: MF_00482; -.
CC InterPro: IPR006244; PsaB
CC InterPro: IPR001280; PSL_PsaB/B.
CC Pfam: PF00223; psaa_psaab; 1.
CC PRINTS: PR00257; PHOTOSYSAAB.
CC TIGRFAMs: TIGR01336; psab; 1.
CC PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB; 1.
CC Photosynthesis; Photosystem I; Transport; Electron transport;
CC Thylakoid; Transmembrane; Metal-binding; Iron; Iron-sulfur; 4Fe-4S;
CC Chlorophyll; Complete proteome.
CC INIT_MET 0 0 BY SIMILARITY.
CC TRANSMEM 45 68 I (POTENTIAL).
CC TRANSMEM 134 157 II (POTENTIAL).
CC TRANSMEM 174 198 III (POTENTIAL).
CC TRANSMEM 272 290 IV (POTENTIAL).
CC TRANSMEM 333 356 V (POTENTIAL).
CC TRANSMEM 372 398 VI (POTENTIAL).
CC TRANSMEM 442 442 VII (POTENTIAL).
CC TRANSMEM 523 541 VIII (POTENTIAL).

Query Match 35.9%; Score 55; DB 1; Length 741;
Best Local Similarity 42.9%; Pred. No. 3.7;
Matches 9; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LICYKESVDQKGNQIMSDKRN 21
      ||| |||::: : : : : ||| |||
      538 LICYKALDARGTKLMPDKD 558
Db 538 LICYKALDARGTKLMPDKD 558

RESULT 9
G3P_MYCPN STANDARD; PRT; 337 AA.
ID G3P_MYCPN
AC P75158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAPA OR GAP OR MPN430 OR MP411.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RA "Complete sequence analysis of the genome of the bacterium Mycoplasma
RA pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000040; AAB96059.1; -
CC PIR: S73737; S73737.
CC HSSP: PL17721; IHG.
CC InterPro: IPR000173; GAP_dhhdhgenase.
CC InterPro: IPR006424; GAPDH-I.
CC Pfam: PF00004; gpdh; 1.
CC Pfam: PF02800; gpdh_C; 1.
CC PRINTS: PR00078; G3PDHGRNASE.
CC TIGRFAMs: TIGR01534; GAPDH-I; 1.
CC PROSITE: PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE
CC ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS

```

```

Immunodominant envelope protein p35 (Virion envelope protein p35).
H3L OR I3L.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
ON NCBI_TaxID=10255;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=India-1967 / Isolate Ind3;
RC MEDLINE=93190624; PubMed=8383392;
RA Shchelkunov S.N., Blinov V.M., Totmenin A.V., Marennikova S.S.,
RA Kolykhalov A.A., Frolov I.V., Chizhikov V.E., Gytarov V.V.,
RA Gashnikov P.V., Belanov E.F., Belayin P.A., Resenchuk S.M.,
RA Andzhaparidze O.G., Sandakchiev L.S.;
RA "Nucleotide sequence analysis of variola virus HindIII M, L, I genome
RT fragments.";
RT Virus Res. 27:25-35(1993).
RL [2]
RP COMPLETE GENOME.
RC STRAIN=India-1967 / Isolate Ind3;
RC MEDLINE=93202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakchiev L.S.;
RA "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RT FEBS Lett. 319:80-83(1993).
RL [3]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
RL [4]
RC EMBL; X67119; CAA47585.1; -
DR EMBL; S55844; AAB24682.1; -
DR EMBL; X69198; CAA49027.1; -
DR PIR; S33100; S33100.
DR InterPro; IPR004900; Pox_P35.
DR Pfam; PF03213; Pox_P35; 1.
DR Envelope protein; Late protein.
KW SEQUENCE 325 AA; 37695 MW; EAB2FEEB3CDEA6BA0 CRC64;
SQ
Query Match 32.7%; Score 50; DB 1; Length 325;
Best Local Similarity 50.0%; Pred. No. 8.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY QY 9 DQKNGQIMSDKRNVIIL 24
| | | | | | | | | |
DB DB 37 DYKDNVEMPEKRNVI 52
| | | | | | | | | |
RESULT 12
VP35_VACC
ID VP35_VACC STANDARD; PRT; 324 AA.
IC P20457;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immunodominant envelope protein p35 (Virion envelope protein p35).
OS H3L.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
ON NCBI_TaxID=10249;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolietti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RL [2]

```

RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.; 'The complete DNA sequence of vaccinia virus'.
RT "Appendix to: 'The complete DNA sequence of vaccinia virus'.
RL Virology 179:517-563(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M35027; AAA48090.1; -
DR PIR: C42514; C42514; Pox_P35.
DR InterPro: IPR004900; Pox_P35.
DR Pfam: PF03213; Pox_P35; 1.
KW Envelope protein; Late protein.
SQ SEQUENCE 324 AA; 37458 MW; ED326A2B3875EB32 CRC64;

Query Match 32.0%; Score 49; DB 1; Length 324;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 9 DQKGNQIMSKRNVIL 24
| | | | | : | | | | :
| | | | | : | | | | :
DQ 36 DVKDNEVMPKRNVVV 51
| | | | | : | | | | :
| | | | | : | | | | :
RESULT 13
VP35_VACCV STANDARD; PRT; 324 AA.
AC P07240;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immunodominant envelope protein p35 (Virion envelope protein p35).
GN H3L.
OS Vaccinia virus (strain WR).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87036903; PubMed=3021979;
RA Rosel J.L., Earl P.L., Weir J.P., Moss B.;
RT "Conserved TAAATG sequence at the transcriptional and translational
RT initiation sites of vaccinia virus late genes deduced by structural
RT and functional analysis of the HindIII H genome fragment."
RL J. Virol. 60:436-449(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13209; AAB59838.1; -
DR PIR: C24481; QOVZH3.
DR InterPro: IPR004900; Pox_P35.
DR Pfam: PF03213; Pox_P35; 1.
KW Envelope protein; Late protein.
SQ SEQUENCE 324 AA; 37504 MW; FF3734631040617F CRC64;

Query Match 32.0%; Score 49; DB 1; Length 324;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 9 DQKGNQIMSKRNVIL 24
| | | | | : | | | | :
| | | | | : | | | | :
DQ 36 DVKDNEVMPKRNVVV 51
| | | | | : | | | | :
| | | | | : | | | | :

Db 36 DVKDNEVMPKRNVVV 51

RESULT 14
GLYD_SCHPO STANDARD; PRT; 467 AA.
AC O13972;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable serine hydroxymethyltransferase, cytosolic (EC 2.1.2.1)
DE (Serine methylase) (Glycine hydroxymethyltransferase) (SHMT).
GN SPAC24C9.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -!- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + glycine +
CC H(2)O = tetrahydrofolate + L-serine.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z98601; CAB11269.1; -
DR PIR: T38353; T38353.
DR HSP; P07511; 1CJ0.
DR GeneDB_SPombe; SPAC24C9.12c; -
DR InterPro: IPR001085; Gly_HyMettransf.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:42 ; Search time 44.75 Seconds
(without alignments)
172.996 Million cell updates/sec

Title: FVIII_ARG581K

Perfect score: 153

Sequence: 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	96.1	2343	6 018806	018806 canis famil
2	141	92.2	2343	6 062730	062730 canis famil
3	82	53.6	2183	11 088783	088783 mus musculus
4	81	52.9	2119	13 090X47	090X47 brachydanio
5	71	46.4	2224	4 043737	043737 homo sapien
6	56	36.6	1157	11 0920H8	0920H8 rattus norv
7	55.5	36.3	1048	6 09X727	09X727 ovis aries
8	54	35.3	847	11 08C4S2	08C4S2 mus musculus
9	54	35.3	1157	11 09Z0Z4	09Z0Z4 mus musculus
10	53.5	35.0	185	16 08D2G5	08D2G5 wigglewort
11	53	34.6	555	2 09RA27	09RA27 vibrio mari
12	52	34.0	207	5 08IL29	08IL29 plasmodium
13	52	34.0	553	5 08IKQ4	08IKQ4 plasmodium
14	51.5	33.7	453	17 0976G7	0976G7 sulfolobus
15	50.5	33.0	357	12 09IGN4	09IGN4 epiphyas po
16	50	32.7	150	16 09A4S1	09A4S1 caulobacter

Q09634 caenorhabdi
Q89217 variola vir
Q85385 variola maj
Q8XMX8 clostridium
Q75180 homo sapien
Q9BQ57 homo sapien
Q9C058 homo sapien
Q855n2 oropouche v
Q8IC67 plasmodium
Q9EMX7 ansacta moo
Q8EXT5 leptospira
Q57206 vaccinia vi
Q8BQ43 mus musculu
Q91F52 arabidopsis
Q10618 helicoverpa
Q91B77 helicoverpa
Q99H34 helicoverpa
Q9RUP8 deinococcus
Q9M078 arabidopsis
Q8239 borrelia bu
Q9S0B2 borrelia bu
Q9AVX2 guillardia
Q8ICY7 plasmodium
Q67249 aquifex aeo
Q75659 homo sapien
Q8SVC4 encephalito
Q8BQK7 mus musculu
Q8BGM9 mus musculu
Q8T235 dictyostell

17 50 32.7 211 5 009634
18 50 32.7 325 12 089217
19 50 32.7 325 12 085385
20 50 32.7 362 16 08XMX8
21 50 32.7 1104 4 075180
22 50 32.7 1158 4 09BQ57
23 50 32.7 1158 4 09C058
24 50 32.7 2250 12 0855N2
25 49.5 32.4 657 5 08IC67
26 49 32.0 157 12 09EMX7
27 49 32.0 179 16 08EXT5
28 49 32.0 324 12 057206
29 49 32.0 355 11 08BQ43
30 49 32.0 641 10 091F52
31 48.5 31.7 183 12 010618
32 48.5 31.7 201 12 091B77
33 48.5 31.7 203 12 099H34
34 48.5 31.7 1078 16 09RUP8
35 48.5 31.7 1817 10 09M078
36 48 31.4 192 2 068239
37 48 31.4 192 2 09S0B2
38 48 31.4 301 10 09AVX2
39 48 31.4 466 5 08ICY7
40 48 31.4 595 16 067249
41 48 31.4 782 4 075659
42 48 31.4 867 5 08SVC4
43 48 31.4 1046 11 08BQK7
44 48 31.4 1162 11 08BGM9
45 48 31.4 1376 5 08T235

ALIGNMENTS

RESULT 1

018806 PRELIMINARY; PRT; 2343 AA.
ID AC 018806;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor VIII.
GN F8.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA Giles A., Lillcrap D.;
RT "The canine factor VIII cDNA and 5' flanking sequence";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL; AF016234; AAB87412.1;
DR HSSP; P00451; LCFG.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR000421; FA58_C
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C.1; 2.
DR PROSITE; PS01286; FA58C.2; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match 96.1%; Score 147; DB 6; Length 2343;
Best Local Similarity 93.3%; Pred. No. 1.1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFDE 30

|||||

```

Db      565 LICYKESVDQGNQMSDKRNVLFSVFDE 594
RESULT 2
O62730
ID      062730      PRELIMINARY;      PRT;      2343 AA.
AC      062730;
DT      01-AUG-1998 (TReMBLrel. 07, Created)
DT      01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Factor VIII.
OS      Canis familiaris (Dog).
OC      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney, and Spleen;
RA      Gordy P.W., Bowen R.A.;
RT      *Characterization of the canine factor VIII cDNA.*;
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR      EMBL; AF049489; AAC05384.1; -.
DR      HSSP; P00451; 1CFG.
DR      InterPro; IPR001117; Cu-oxidase.
DR      InterPro; IPR000421; FA58_C.
DR      Pfam; PF00394; Cu-oxidase; 3.
DR      Pfam; PF00754; F5_F8_type_C; 2.
DR      SMART; SM00231; FA58C; 2.
DR      PROSITE; PS01285; FA58C_1; 2.
DR      PROSITE; PS01286; FA58C_2; 2.
DR      PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ      SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match      92.2%; Score 141; DB 6; Length 2343;
Best Local Similarity 90.0%; Pred. No. 8.3e-12;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 LICYKESVDQGNQMSDKRNVLFSVFDE 30
      ||||| ||||| ||||| ||||| ||||| |||||
Db      565 LICYKESVDQGNQMSDKRNVLFSVLDE 594

RESULT 3
O88783
ID      088783      PRELIMINARY;      PRT;      2183 AA.
AC      088783;
DT      01-NOV-1998 (TReMBLrel. 08, Created)
DT      01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Murine coagulation factor V.
GN      F5.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=98282202; PubMed=9616155;
RA      Yang T.L., Cui J., Rehumtulla A., Yang A., Moussalli M., Kaufman R.J.,
RA      Ginsburg D.;
RT      *The structure and function of murine factor V and its inactivation by
RT      protein C.*;
RL      Blood 91:4593-4598(1998).
CC      -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR      EMBL; U52925; AAC99553.1; -.
DR      HSSP; P12259; 1CZT.
DR      MGD; MGI:88382; F5.
DR      InterPro; IPR001117; Cu-oxidase.
DR      InterPro; IPR000421; FA58_C.
DR      Pfam; PF00394; Cu-oxidase; 3.
DR      Pfam; PF00754; F5_F8_type_C; 2.
DR      SMART; SM00231; FA58C; 2.

```

```

DR      PROSITE; PS01285; FA58C_1; 2.
DR      PROSITE; PS01286; FA58C_2; 2.
DR      PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ      SEQUENCE 2183 AA; 247228 MW; BF0A8AA723F60317 CRC64;

Query Match      53.6%; Score 82; DB 11; Length 2183;
Best Local Similarity 53.3%; Pred. No. 0.0047;
Matches 16; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY      1 LICYKESVDQGNQMSDKRNVLFSVFDE 30
      ||||| : : : : : ||||| : : : : : |||||
Db      190 LICKKGTLEDGTQKMPDKOHVLLFAVFDE 219

RESULT 4
Q90X47
ID      Q90X47      PRELIMINARY;      PRT;      2119 AA.
AC      Q90X47;
DT      01-DEC-2001 (TReMBLrel. 19, Created)
DT      01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      SC:B2015.3 (Novel protein similar to vertebrate coagulation factor V
DE      and VIII).
DE      SC:B2015.3.
GN      Brachydanio rerio (Zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Lloyd D.;
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR      EMBL; AL590146; CAC94896.1; -.
DR      InterPro; IPR001117; Cu-oxidase.
DR      Pfam; PF00394; Cu-oxidase; 2.
DR      Pfam; PF00754; F5_F8_type_C; 2.
DR      SMART; SM00231; FA58C; 2.
DR      PROSITE; PS01285; FA58C_1; 2.
DR      PROSITE; PS01286; FA58C_2; 1.
DR      PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ      SEQUENCE 2119 AA; 240643 MW; DC0E806FFA8761E6 CRC64;

Query Match      52.9%; Score 81; DB 13; Length 2119;
Best Local Similarity 53.3%; Pred. No. 0.0064;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 LICYKESVDQGNQMSDKRNVLFSVFDE 30
      ||||| : : : ||||| : : : |||||
Db      182 LICKKGTLEDGSGNQIHQESVLLFGVFDE 211

RESULT 5
O43737
ID      O43737      PRELIMINARY;      PRT;      2224 AA.
AC      O43737;
DT      01-JUN-1998 (TReMBLrel. 06, Created)
DT      01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Factor V.
DE      Homo sapiens (Human).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Bird C.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR      EMBL; Z99572; CAB16748.1; -.
DR      HSSP; P12259; 1CZT.

```



```

RA Gitschier J., Anderson G.J.;
RT "Hephaestin, a ceruloplasmin homologue implicated in intestinal iron
RT transport, is defective in the sla mouse.";
RL Nat. Genet. 21:195-199(1999).
DR EMBL; AF082567; AAD16035.1; -
DR HSSP; P00450; 1KCW.
DR MGD; MGI:1332240; Heph.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; MultiCu_oxidse2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1157 AA; 129682 MW; 411BBIDA2BFEC2FD CRC64;

Query Match 35.3%; Score 54; DB 11; Length 1157;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 12; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFE 30
| : | : | : | : | : | : | : | : |
Db 557 LVCKAGALGADGKQKGVKREFFLLTFVFE 586

RESULT 10
Q8D2G5 PRELIMINARY; PRT; 185 AA.
AC Q8D2G5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Frr protein.
GN FRR.
OS Wigglesworthia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=164609;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia.";
RL Nat. Genet. 32:402-407(2002).
DR EMBL; AB063522; BAC24535.1; -
KW Complete proteome.
SQ SEQUENCE 185 AA; 21650 MW; C71569E7155DECFD CRC64;

Query Match 35.0%; Score 53.5; DB 16; Length 185;
Best Local Similarity 42.9%; Pred. No. 5.8;
Matches 12; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 3 CYKESVD-QKGNQIMSDKRNVLFSVFD 29
| | | | | : | : | : | : | : |
Db 44 CYGTSVPLSKLSNLSEKSNILKINVPD 71

RESULT 11
Q9RA27 PRELIMINARY; PRT; 555 AA.
AC Q9RA27;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Genes, similar to eicosapentaenoic acid synthesis gene cluster.
OS Vibrio marinus (Moritella marina).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Moritella.
OX NCBI_TaxID=90736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MP-1;
RA Tanaka M., Ueno A., Kawasaki K., Yumoto I., Ohgiya S., Hoshino T.,

```

```

RA Ishizaki K., Okuyama H., Morita N.;
RT "Isolation of clustered genes that are notably homologous to the
RT eicosapentaenoic acid biosynthesis gene cluster from the
RT docosahexaenoic acid-producing bacterium Vibrio marinus strain MP-1.";
RL Biotechnol. Lett. 21:939-945(1999).
DR EMBL; AB025342; BAA89376.1; -
DR HSSP; P22106; 1CT9.
DR InterPro; IPR001962; Asn_synthase.
DR InterPro; IPR006426; Asn_synth_AEB.
DR InterPro; IPR000583; GATase_2.
DR Pfam; PF00733; Asn_synthase; 1.
DR Pfam; PF00310; GATase_2; 1.
DR TIGRFAMS; TIGR01536; asn_synth_AEB; 1.
DR PROSITE; PS00443; GATASE_TYPE-II; 1.
SQ SEQUENCE 555 AA; 62544 MW; 4071ECC0990C2G34 CRC64;

Query Match 34.6%; Score 53; DB 2; Length 555;
Best Local Similarity 36.7%; Pred. No. 22;
Matches 11; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY 1 LICYKESVDQKGNQIMSDKRNVLFSVFE 30
| | | | | : | : | : | : | : |
Db 104 LALYKE---KGTQFLDDLNGIFAFALYDE 129

RESULT 12
Q8IL29 PRELIMINARY; PRT; 207 AA.
AC Q8IL29;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pr14_0420.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=3D7;
RC MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014823; AAN37033.1; -
KW Hypothetical protein.
SQ SEQUENCE 207 AA; 24953 MW; 58C249B4C5A46214 CRC64;

Query Match 34.0%; Score 52; DB 5; Length 207;
Best Local Similarity 42.3%; Pred. No. 11;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 4 YKESVDQKGNQIMSDKRNVLFSVFD 29
| | | | | : | : | : | : | : |
Db 108 YLRVNDKKKQVPSDLYNIQFFMFD 133

RESULT 13
Q8IKQ4 PRELIMINARY; PRT; 553 AA.
AC Q8IKQ4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

```

```

RESULT 15
Q91GN4 PRELIMINARY; PRT; 357 AA.
ID Q91GN4;
AC Q91GN4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 41.4 kDa protein.
OS Epiphyas postvittana nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
CX NCBI_TaxID=70600;
RX [1]
RN SEQUENCE FROM N.A.
RP RA Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., I.
RA Ward V.K.;
RT "The complete sequence of the Epiphyas postvittana
RT rt nucleopolyhedrovirus genome.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY043265; AAK85579.1; -.
KW Hypothetical protein.
SQ SEQUENCE 357 AA; 1384 MW; 9EAE830B66132408 CRC64;
Query Match 33.0%; Score 50.5; DB 12; Length
Best Local Similarity 40.0%; Pred. No. 33;
Matches 10; Conservative 8; Mismatches 6; Indels
Qy 5 KESVDQKGNQ-IMGDKRNVILFSVF 28
Db 171 QNGVDEYNNQKIMINKRHAVTFNVF 195

```

Search completed: October 21, 2003, 18:57:29
Job time : 47.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:44:27 ; Search time 49.5 Seconds
(without alignments)
96.198 Million cell updates/sec

Title: FVIII_PHE328S

Perfect score: 158
Sequence: 1 TFLTAQTLLMDLGGFLLSCHISSHQHDSME 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	100.0	1383	18 AAW33229	Procoagulant-activ
2	158	100.0	2332	18 AAW33225	Procoagulant-activ
3	158	100.0	2332	18 AAW33226	Procoagulant-activ
4	152	96.2	372	16 AAR73019	Human Factor-VIII
5	152	96.2	400	16 AAR67709	B-domain deleted F
6	152	96.2	720	16 AAR74088	Factor-VIII heavy
7	152	96.2	729	16 AAR74089	Factor-VIII heavy
8	152	96.2	740	16 AAR76959	Human Factor-VIII/
9	152	96.2	740	16 AAR76961	Human Factor-VIII

10	152	96.2	740	16 AAR76962	Human Factor-VIII
11	152	96.2	740	16 AAR73021	Human Factor-VIII
12	152	96.2	740	16 AAR74090	Factor-VIII heavy
13	152	96.2	740	16 AAR76982	Human factor VIII
14	152	96.2	1014	8 AAF71139	Factor VIII:c varia
15	152	96.2	1383	18 AAW33227	Procoagulant-activ
16	152	96.2	1383	18 AAW33228	Procoagulant-activ
17	152	96.2	1424	9 AAP80268	Modified factor VI
18	152	96.2	1424	10 AAP91169	Sequence of 740 Ar
19	152	96.2	1424	22 AAB48842	Mutant mature huma
20	152	96.2	1424	23 AAO18622	Human mature B-dom
21	152	96.2	1425	9 AAP80267	Modified factor VI
22	152	96.2	1438	21 AAB01262	B-domain deleted f
23	152	96.2	1440	12 AAR12971	Factor VIII:SQ. U
24	152	96.2	1445	23 ABC92540	LE B-domain-dele
25	152	96.2	1447	23 ABC92541	5Arg B-domain-dele
26	152	96.2	1457	19 AAW46246	Human factor VIII
27	152	96.2	1457	19 AAW44372	Human Factor VIII
28	152	96.2	1457	20 AAY21675	Beta-domain delete
29	152	96.2	1459	22 AAE10827	Human factor VIII
30	152	96.2	1459	22 AAE10832	Human factor VIII
31	152	96.2	1459	22 AAE10833	Human factor VIII
32	152	96.2	1471	18 AAW23414	Human B-domain del
33	152	96.2	1471	22 AAE67959	Amino acid sequenc
34	152	96.2	1516	9 AAP80265	Modified factor VI
35	152	96.2	1661	18 AAW18670	Factor VIII-dH695-
36	152	96.2	2098	17 AAR6863	Factor-VIII. Homo
37	152	96.2	2332	8 AAP71726	Factor VIII:c varia
38	152	96.2	2332	8 AAP71727	Factor VIII:c varia
39	152	96.2	2332	8 AAP71728	Factor VIII:c varia
40	152	96.2	2332	8 AAP71729	Factor VIII:c varia
41	152	96.2	2332	14 AAR43257	Human Factor VIII.
42	152	96.2	2332	18 AAW33222	Procoagulant-activ
43	152	96.2	2332	18 AAW33223	Procoagulant-activ
44	152	96.2	2332	18 AAW33224	Procoagulant-activ
45	152	96.2	2332	19 AAW53483	Human factor VIII.

ALIGNMENTS

RESULT 1

AAW33229
ID AAW33229 standard; protein; 1383 AA.

AC AAW33229;

DT 30-APR-1998 (first entry)

DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.

KW Pro-coagulant active factor VIII; FVIII; haemophilia A;

KW recombinant secretion; pro-coagulant activity; resistance;

KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;

KW von Willebrand factor binding site; binding affinity;

KW FVIII replacement therapy.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..346

FT Region /note= "factor VIIIA heavy chain"

FT Region 741..1383

FT Region /note= "factor VIIIA light chain"

FT Domain 1..329

FT Domain /note= "A1 domain"

FT Domain 1..179

FT Domain /note= "plastocyanin-like domain 1"

FT Domain 187..329

FT Domain /note= "plastocyanin-like domain 2"

FT Domain 380..711

FT Domain /note= "A2 domain"

FT Misc_feature 711..746
 FT /note= "a spacer of the sequence
 FT SFSONSRHPSTROKOFNATIPENDIEKTDWPF
 FT AHRTMPKIQNVSSDLMLL is inserted
 FT between domains A2 and A3"
 FT
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 746..1073
 FT /note= "A3 domain"
 FT 1073..1221
 FT /note= "C1 domain"
 FT Domain 1226..1378
 FT /note= "C2 domain"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Misc-difference 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 FT
 FT W09740145-A1.
 FT
 FT 30-OCT-1997.
 FT
 FT 24-APR-1997; 97W0-US06563.
 FT
 FT 15-MAY-1996; 96US-0017785.
 FT 24-APR-1996; 96US-0016117.
 FT
 FT (UNWI) UNIV MICHIGAN.
 FT
 FT Amano K, Kaufman RJ, Pipe SW;
 FT WPI; 1997-535830/49.
 FT
 FT Modified human pro-coagulant active factor VIII - can be
 FT administered to haemophiliacs, i.e. factor VIII replacement therapy
 FT
 FT Claim 19; Page -; 57pp; English.
 FT
 FT The present sequence represents a novel pro-coagulant active factor
 FT VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 FT von Willebrand factor binding site, mutations F309S, R740A and addition
 FT of an amino acid sequence spacer between the A2 and A3 domains. Factor
 FT VIII, along with calcium and phospholipid, acts as a cofactor for factor
 FT IXA, when it converts factor X to the activated form (factor XA). FVIII
 FT is the coagulation factor deficient in the X-chromosome-linked bleeding
 FT disorder haemophilia A. Several other mutant FVIII proteins have also
 FT been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
 FT capable of recombinant secretion at higher levels than typically obtained
 FT with wild type FVIII and retains pro-coagulant activity. The FVIII
 FT mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
 FT protein C (APC) cleavage. The present FVIII mutant can form a more
 FT stable configuration, and have an approximate 5-fold increase in
 FT specific activity compared to purified wild type FVIII, while increasing
 FT their binding affinity to von Willebrand factor improves their
 FT stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 FT FVIII replacement therapy, while the nucleic acid molecule can be used
 FT for gene therapy.
 FT note: this sequence does not appear in the specification; it was created
 FT using sequences from the given references.
 FT
 FT SQ Sequence 1383 AA;
 FT
 FT Query Match 100.0%; Score 158; DB 18; Length 1383;

Best Local Similarity 100.0%; Pred. No. 1.2e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TFLTAQTLLMDLQGLSQFLSSCHSHQHDGME 30
 DB 292 TFLTAQTLLMDLQGLSQFLSSCHSHQHDGME 321
 RESULT 2
 AAW33225
 ID AAW33225 standard; protein; 2332 AA.
 XX AC AAW33225;
 XX DT 30-APR-1998 (first entry)
 XX DE Procoagulant-active human factor VIII:C (FVIII) F309S mutant protein.
 XX KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX OS Synthetic.
 OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Region 1..1313
 FT /note= "factor VIIIA heavy chain"
 FT Region 1649..2332
 FT /note= "factor VIIIA light chain"
 FT Domain 1..329
 FT /note= "A1 domain"
 FT Domain 1..179
 FT /note= "plastocyanin-like domain 1"
 FT Domain 187..329
 FT /note= "plastocyanin-like domain 2"
 FT Domain 380..711
 FT /note= "A2 domain"
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 741..1648
 FT /note= "B domain"
 FT Domain 1694..2021
 FT /note= "A3 domain"
 FT Domain 1694..1858
 FT /note= "plastocyanin-like domain 5"
 FT Domain 1888..2021
 FT /note= "plastocyanin-like domain 6"
 FT Domain 2021..2169
 FT /note= "C1 domain"
 FT Domain 2174..2325
 FT /note= "C2 domain"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Cleavage-site 740..741
 FT /note= "by thrombin"
 FT Cleavage-site 1689..1690
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Disulfide-bond 1832..1858
 FT /note= "probable"
 FT Misc-difference 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT
 FT W09740145-A1.

XX 30-OCT-1997.
PD 30-OCT-1997.
XX 24-APR-1997; 97WO-US06563.
XX 15-MAY-1996; 96US-0017785.
PR 24-APR-1996; 96US-0016117.
XX (UNMI) UNIV MICHIGAN.
XX Amano K, Kaufman RJ, Pipe SW;
XX WPI; 1997-535830/49.
XX Modified human pro-coagulant active factor VIII - can be
PT administered to haemophiliacs, i.e. factor VIII replacement therapy
XX Claim 4; Page -: 57pp; English.
XX The present sequence represents a novel pro-coagulant active factor
CC VIII (FVIII) mutant protein, comprising a mutation at Phe309, which is
CC replaced with Ser. Factor VIII, along with calcium and phospholipid,
CC acts as a cofactor for factor IXA, when it converts factor X to the
CC activated form (factor IXA). FVIII is the coagulation factor deficient in
CC the X-chromosome-linked bleeding disorder haemophilia A. Several other
CC mutant FVIII proteins have also been created (see AAW33222-29). The
CC FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
CC higher levels than typically obtained with wild type FVIII and retains
CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
CC (AAW33223) are resistant to activated protein C (APC) cleavage. The
CC FVIII mutant comprising a deletion of the B domain and von Willebrand
CC factor binding site, a mutation at Arg740 and an addition of an amino
CC acid sequence spacer between the A2 and A3 domains can form a more
CC stable configuration, and have an approximate 5-fold increase in
CC specific activity compared to purified wild type FVIII, while increasing
CC their binding affinity to von Willebrand factor improves their
CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
CC FVIII replacement therapy, while the nucleic acid molecule can be used
CC for gene therapy.
CC note: this sequence does not appear in the specification; it was created
CC using sequences from the given references.
XX SQ Sequence 2332 AA;
Query Match 100.0%; Score 158; DB 18; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TELTAQTLLMDLGQFLLSCHSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLLSCHSHQHDGME 321
RESULT 3
AAW33226
ID AAW33226 standard; protein; 2332 AA.
XX AC AAW33226;
XX DT 30-APR-1998 (first entry)
XX Procoagulant-active factor VIII:C F309S/R336I/R562K mutant protein.
XX Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers

FT Region 1..1313 "factor VIIIA heavy chain"
FT /note= "2332"
FT Region 1649..2332
FT /note= "factor VIIIA light chain"
FT Domain 1..329
FT /note= "A1 domain"
FT Domain 1..179
FT /note= "plastocyanin-like domain 1"
FT Domain 187..329
FT /note= "plastocyanin-like domain 2"
FT Domain 380..711
FT /note= "A2 domain"
FT Domain 380..554
FT /note= "plastocyanin-like domain 3"
FT Domain 564..711
FT /note= "plastocyanin-like domain 4"
FT Domain 741..1648
FT /note= "B domain"
FT Domain 1694..2021
FT /note= "A3 domain"
FT Domain 1694..1858
FT /note= "plastocyanin-like domain 5"
FT Domain 1868..2021
FT /note= "plastocyanin-like domain 6"
FT Domain 2021..2169
FT /note= "C1 domain"
FT Domain 2174..2326
FT /note= "C2 domain"
FT Cleavage-site 372..373
FT /note= "by thrombin"
FT Cleavage-site 740..741
FT /note= "by thrombin"
FT Cleavage-site 1689..1690
FT /note= "by thrombin"
FT Disulfide-bond 153..179
FT /note= "probable"
FT Disulfide-bond 528..554
FT /note= "probable"
FT Disulfide-bond 1832..1858
FT /note= "probable"
FT Misc-difference 309
FT /label= F309S
FT /note= "wild type Phe replaced with Ser"
FT Misc-difference 336
FT /label= R336I
FT /note= "wild type Arg replaced with Ile"
FT Misc-difference 562
FT /label= R562K
FT /note= "wild type Arg replaced with Lys"
XX WO9740145-A1.
XX 30-OCT-1997.
XX 24-APR-1997; 97WO-US06563.
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96US-0016117.
XX (UNMI) UNIV MICHIGAN.
XX Amano K, Kaufman RJ, Pipe SW;
XX WPI; 1997-535830/49.
XX Modified human pro-coagulant active factor VIII - can be
PT administered to haemophiliacs, i.e. factor VIII replacement therapy
XX Claim 11; Page -: 57pp; English.
XX The present sequence represents a novel pro-coagulant active factor
CC VIII (FVIII) mutant protein, comprising the mutations Phe309Ser,
CC Arg336Ile, and Arg562Lys. Factor VIII, with calcium and phospholipid,

CC acts as a cofactor for factor IXA, when it converts factor X to the
 CC activated form (factor XA). FVIII is the coagulation factor deficient in
 CC the X-chromosome-linked bleeding disorder haemophilia A. Several other
 CC mutant FVIII proteins have also been created (see AAW33222-29). The
 CC FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
 CC higher levels than typically obtained with wild type FVIII and retains
 CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R526K
 CC (AAW33223) are resistant to activated protein C (APC) cleavage. The
 CC FVIII mutant comprising a deletion of the B domain and von Willebrand
 CC factor binding site, a mutation at Arg740 and an addition of an amino
 CC acid sequence spacer between the A2 and A3 domains can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

XX SQ Sequence 2332 AA;
 Query Match 100.0%; Score 158; DB 18; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 2.2e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 30
 |||||
 Db 292 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 321

RESULT 4
 AAR73019
 ID AAR73019 standard; peptide; 372 AA.

XX AC AAR73019;
 XX XX
 DT 25-MAR-2003 (updated)
 DT 21-NOV-1995 (first entry)
 XX XX
 DE Human Factor-VIII N-terminal fragment.

XX KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.

XX OS Homo sapiens.

XX PN WO9513301-A1.

XX PD 18-MAY-1995.

XX PF 10-NOV-1994; 94WO-DK00424.

XX PR 12-NOV-1993; 93DK-0001281.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Persson E;

XX DR WPI; 1995-194038/25.

XX CC Crosslinked Factor VIII polypeptide which is stable - is prepd. using
 PT bis(sulphosuccinimidy) suberate or disuccinimidy suberate in the
 PT presence of polysorbate 80 to produce a coagulant with long lasting
 PT activity

XX PS Disclosure; Page 18; 36pp; English.

XX CC This is the N-terminal fragment of human Factor-VIII which

CC may be crosslinked resulting in increased stability and retention
 CC of high activity over extended periods of time after activation by
 CC thrombin. The polypeptide is used to prevent or treat diseases
 CC caused by the absence or deficiency of Factor-VIII in a subject
 CC such as haemophilia.

CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 372 AA;

Query Match 96.2%; Score 152; DB 16; Length 372;
 Best Local Similarity 96.7%; Pred. No. 2.2e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 30
 |||||
 Db 292 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 321

RESULT 5
 AAR67709
 ID AAR67709 standard; Protein; 400 AA.

XX AC AAR67709;

XX XX

DT 25-MAR-2003 (updated)

DT 20-JUL-1995 (first entry)

XX XX B-domain deleted Factor-VIII.

DE Factor-VIII; blood-clotting; hemophilia A; gene therapy;

KW adenovirus; vector.

XX OS Homo sapiens.

XX PN WO9429471-A1.

XX XX 22-DEC-1994.

XX PF 13-APR-1994; 94WO-US04075.

XX XX 10-JUN-1993; 93US-0074920.

PR 25-MAR-1994; 94US-0218335.

XX XX (GENE-) GENETIC THERAPY INC.

XX PI Connelly S, Kaleko M, Smith T;

XX DR WPI; 1995-036495/05.

DR N-PSDB; AAQ76016.

XX XX New adenoviral vectors for treatment of haemophilia - contg. a
 PT DNA sequence encoding a clotting factor, partic. Factor VIII or
 PT Factor IX

XX PS Disclosure; Fig. 17A-17C; 116pp; English.

XX CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
 CC used to construct recombinant adenovirus vectors that produced
 CC therapeutic levels of the clotting factor when administered to an
 CC animal host, potentially providing hemophilia A gene therapy.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 400 AA;

Query Match 96.2%; Score 152; DB 16; Length 400;
 Best Local Similarity 96.7%; Pred. No. 2.4e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 30
 |||||
 Db 310 TFLTAQTLLMDLQGLFLLSCHSHQHDGME 339

RESULT 6
 AAR74088
 ID AAR74088 standard; protein; 720 AA.

XX XX
 AC AAR74088;


```

FT      /label= Glu absent or Gln, Ser, Thr, Ala
FT      Misc-difference 440
FT      /label= Glu absent or Gln, Ser, Thr, Ala
XX
XX      WO9518829-A1.
XX
XX      13-JUL-1995.
XX
XX      06-JAN-1995; 95WO-DK00010.
XX
XX      07-JAN-1994; 94DK-0000030.
XX
XX      (NOVO ) NOVO-NORDISK AS.
XX
XX      Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX      WPI; 1995-255041/33.
XX
XX      Novel factor VIII derivative used to treat haemophilia - and
XX      comprises a functional A2 domain containing a mutation at one or
XX      more Cys residues.
XX
XX      Disclosure; Page 11-14; 24pp; English.
XX
XX      The new Factor-VIII/Factor-VIIa derivative comprises a functional
XX      A2 domain in which amino acid residues, as indicated in the
XX      features, are deleted or substituted by another amino acid so as to
XX      increase the overall positive charge. Asp-403, Asp-404, Asp-433,
XX      Asp-482 and Asp-500 are preferably substituted by Asn; Gln-434 and
XX      Gln-440 are preferably substituted by Gln. The new derivative has
XX      the same activity as the wild-type Factor-VIII but with improved
XX      stability (the activity is maintained for a longer period compared
XX      to the rapid decline of the activity of wt Factor-VIII). The new
XX      derivative can be used in a composition for treating diseases caused
XX      by an absence or deficiency of Factor-VIII, especially haemophilia.
XX      (Updated on 25-MAR-2003 to correct DR field.)
XX
XX      Sequence 740 AA;
XX
XX      Query Match 96.2%; Score 152; DB 16; Length 740;
XX      Best Local Similarity 96.7%; Pred. No. 4.8e-14;
XX      Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX      1 TFLTAQTLLMDLQGFLLSCHSHQHDGME 30
XX      |||||
XX      292 TFLTAQTLLMDLQGFLLSCHSHQHDGME 321
XX
XX      RESULT 9
XX      AAR76961
XX      ID AAR76961 standard; protein; 740 AA.
XX
XX      AC AAR76961;
XX
XX      09-MAR-1996 (first entry)
XX
XX      Human Factor-VIII derivative.
XX
XX      Factor-VIII; therapeutic; blood-clotting.
XX
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      Misc-difference 692 /label= absent or Ala, Thr, Ser, Gly or Asp
XX      Misc-difference 720 /label= absent or Gln, Ser, Thr, Val or Ala
XX      Misc-difference 729 /label= absent or Val, Ala or Ile
XX
XX      WO9518827-A1.
XX
XX      13-JUL-1995.
XX
XX      06-JAN-1995; 95WO-DK00008.
XX
XX      07-JAN-1994; 94DK-0000032.
XX
XX      (NOVO ) NOVO-NORDISK AS.
XX
XX      Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX      13-JUL-1995.

```

```

XX      06-JAN-1995; 95WO-DK00008.
XX
XX      07-JAN-1994; 94DK-0000032.
XX
XX      (NOVO ) NOVO-NORDISK AS.
XX
XX      Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX      WPI; 1995-255039/33.
XX
XX      Novel factor VIII derivative used to treat haemophilia - and
XX      comprises a functional A2 domain containing a mutation at one or
XX      more Cys residues.
XX
XX      Disclosure; Page 14-17; 30pp; English.
XX
XX      The new Factor-VIII derivative comprises a functional A2 domain in
XX      which Cys-692 is deleted or replaced with another amino acid
XX      residue, preferably Ser (see AAR76962). Alternatively, Glu-720 and/or
XX      Tyr-729 are deleted or substituted with various amino acids (as in
XX      the features). The new derivative has the same activity as the wild-
XX      type Factor-VIII but with improved stability (the activity is
XX      maintained for a longer period compared to the rapid decline of the
XX      activity of wt Factor-VIII). The new derivative can be used in a
XX      composition for treating diseases caused by an absence or deficiency
XX      of Factor-VIII, especially haemophilia.
XX
XX      Sequence 740 AA;
XX
XX      Query Match 96.2%; Score 152; DB 16; Length 740;
XX      Best Local Similarity 96.7%; Pred. No. 4.8e-14;
XX      Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX      1 TFLTAQTLLMDLQGFLLSCHSHQHDGME 30
XX      |||||
XX      292 TFLTAQTLLMDLQGFLLSCHSHQHDGME 321
XX
XX      RESULT 10
XX      AAR76962
XX      ID AAR76962 standard; protein; 740 AA.
XX
XX      AC AAR76962;
XX
XX      09-MAR-1996 (first entry)
XX
XX      Human Factor-VIII derivative.
XX
XX      Factor-VIII; therapeutic; blood-clotting.
XX
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      Misc-difference 692 /label= Cys substituted by Ser
XX      Misc-difference 720 /label= absent or Gln, Ser, Thr, Val or Ala
XX      Misc-difference 729 /label= absent or Val, Ala or Ile
XX
XX      WO9518827-A1.
XX
XX      13-JUL-1995.
XX
XX      06-JAN-1995; 95WO-DK00008.
XX
XX      07-JAN-1994; 94DK-0000032.
XX
XX      (NOVO ) NOVO-NORDISK AS.
XX
XX      Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX      13-JUL-1995.

```

DR WPI; 1995-255039/33.
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX
 XX Disclosure; Page 18-20; 30pp; English.
 XX
 CC The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is replaced with Ser. For other (less preferred)
 CC substitutions at this site, see AAR76961. Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX
 XX Sequence 740 AA;
 SQ
 Query Match 96.2%; Score 152; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 4.8e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
 Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
 RESULT 11
 AAR73021
 ID AAR73021 standard; peptide; 740 AA.
 XX
 XX AAR73021;
 XX
 XX 25-MAR-2003 (updated)
 DT 21-NOV-1995 (first entry)
 DT
 XX Human Factor-VIII N-terminal fragment.
 DE
 XX Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
 KW
 XX Homo sapiens.
 OS
 XX WO9513301-A1.
 PN
 XX 18-MAY-1995.
 PD
 XX 10-NOV-1994; 94WO-DK00424.
 PF
 XX 12-NOV-1993; 93DK-0001281.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Persson E;
 PI
 XX WPI; 1995-194038/25.
 DR
 XX Crosslinked Factor VIII polypeptide which is stable - is prepd. using
 PT bis(sulphosuccinimidy)l suberate or disuccinimidy)l suberate in the
 PT presence of polysorbate 80 to produce a coagulant with long lasting
 PT activity
 XX
 XX Disclosure; Page 21; 36pp; English.
 PS
 XX This is the N-terminal fragment of human Factor-VIII which may be
 CC crosslinked resulting in increased stability and retention of high
 CC activity over extended periods of time after activation by thrombin.
 CC The polypeptide is used to prevent or treat diseases caused by the
 CC absence or deficiency of Factor-VIII in a subject such as
 CC haemophilia.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX
 SQ Sequence 740 AA;
 XX
 Query Match 96.2%; Score 152; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 4.8e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
 Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
 RESULT 12
 AAR74090
 ID AAR74090 standard; protein; 740 AA.
 XX
 XX AAR74090;
 XX
 XX 25-MAR-2003 (updated)
 DT 04-NOV-1995 (first entry)
 DT
 XX Factor-VIII heavy chain N-terminal fragment.
 DE
 XX human; Factor VIII; heavy chain; N-terminal fragment;
 KW thrombin cleavage; blood-clotting.
 KW
 XX Homo sapiens.
 OS
 XX WO9513300-A1.
 PN
 XX 18-MAY-1995.
 PD
 XX 10-NOV-1994; 94WO-DK00423.
 PF
 XX 12-NOV-1993; 93DK-0001280.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Ezban Rasmussen M, Kjalke M;
 PI
 XX WPI; 1995-194037/25.
 DR
 XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
 PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
 PT activity, may be used to treat patients who have developed antibodies to
 PT C-terminal epitope(s) of Factor VIII
 XX
 XX Disclosure; Page 30-32; 51pp; English.
 PS
 XX The sequence represents N-terminal residues 1-740 of a human Factor-
 CC VIII heavy chain. The sequence contains entire A1 and A2 domains,
 CC and truncated forms (1-720 in (AAR74088) and 1-729 in (AAR74090)).
 CC may be produced by treatment with a protease, e.g. thrombin. The
 CC C-terminally truncated fragments have the same coagulant specific
 CC activity as full-length Factor-VIII, and may be produced
 CC recombinantly to reduce production costs and improve safety, giving
 CC higher production levels and stability than for the full-length form.
 CC The fragments may be used to treat patients who have developed
 CC antibodies against epitopes in the C-terminal part of the heavy chain.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 740 AA;
 SQ
 Query Match 96.2%; Score 152; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 4.8e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
 Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
 RESULT 13

ID AAP71139 standard; Protein; 1014 AA.

KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.

XX Synthetic.
 OS Homo sapiens.

XX Key
 FT Region Location/Qualifiers
 FT 1..346 /note= "factor VIIIA heavy chain"
 FT 741..1383 /note= "factor VIIIA light chain"
 FT Domain 1..329 /note= "A1 domain"
 FT Domain 1..179 /note= "plastocyanin-like domain 1"
 FT Domain 187..329 /note= "plastocyanin-like domain 2"
 FT Domain 380..711 /note= "A2 domain"
 FT MISC_feature 711..746 /note= "a spacer of the sequence
 SFSNSRHPSTRQKQFNATIPENDIEKTDWPF
 AHRTPMPKIONVSSDLMLL is inserted
 between domains A2 and A3"
 FT Domain 380..554 /note= "plastocyanin-like domain 3"
 FT Domain 564..711 /note= "plastocyanin-like domain 4"
 FT Domain 746..1073 /note= "A3 domain"
 FT Domain 1073..1221 /note= "C1 domain"
 FT Domain 1226..1378 /note= "C2 domain"
 FT Cleavage-site 372..373 /note= "by thrombin"
 FT Disulfide-bond 153..179 /note= "probable"
 FT Disulfide-bond 528..554 /note= "probable"
 FT Misc-difference 740 /label= R740A
 FT /note= "wild type Arg replaced with Ala"

XX WO9740145-A1.

XX 30-OCT-1997.

XX 24-APR-1997; 97WO-US06563.

XX 15-MAY-1996; 96US-0017785.

XX 24-APR-1996; 96US-0016117.

XX (UNMI) UNIV MICHIGAN.

XX Amano K, Kaufman RJ, Pipe SW;
 PI WPI; 1997-535830/49.

XX Modified human pro-coagulant active factor VIII - can be
 PT administered to haemophiliacs, i.e. factor VIII replacement therapy

XX Claim 20: Page -; 57pp; English.

XX The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 CC von Willebrand factor binding site, a mutation at Arg740 and an addition
 CC of an amino acid sequence spacer between the A2 and A3 domains. Factor
 CC VIII, along with calcium and phospholipid, acts as a cofactor for factor
 CC IXa, when it converts factor X to the activated form (factor Xa). FVIII
 CC is the coagulation factor deficient in the X-chromosome-linked bleeding
 CC disorder haemophilia A. Several other mutant FVIII proteins have also
 CC been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is

CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII
 CC mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
 CC protein C (APC) cleavage. The present FVIII mutant can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

XX Sequence 1383 AA;

Query Match 96.2%; Score 152; DB 18; Length 1383;
 Best Local Similarity 96.7%; Pred. No. 9.8e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQFLMLDGLGQFLLSCHISSHQHDGME 30
 ||||||||||||||||||||||||||||
 Db 292 TFLTAQFLMLDGLGQFLLSCHISSHQHDGME 321

Search completed: October 21, 2003, 18:53:30
 Job time : 50.5 secs

GenCore version 5.1.6
 Copyr:ght (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:50:07 ; Search time 16 Seconds
 (without alignments)
 79.333 Million cell updates/sec

Title: FVIII_PHE328S
 Perfect score: 158
 Sequence: 1 TFLTAQTLLMDLQGLLSCHISSHQHDGME 30

Scoring table: BLOSUM62DX
 Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents AA: *
 1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
 2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
 3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
 4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
 5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	152	96.2	1438	4	US-09-209-916-1
2	152	96.2	1471	1	US-08-683-839B-3
3	152	96.2	1661	2	US-08-882-083-2
4	152	96.2	1661	2	US-08-558-107-2
5	152	96.2	1661	3	US-09-243-539-2
6	152	96.2	2332	1	US-07-864-004B-4
7	152	96.2	2332	1	US-08-251-937A-4
8	152	96.2	2332	1	US-08-212-133A-2
9	152	96.2	2332	1	US-08-276-594A-2
10	152	96.2	2332	1	US-08-474-503-2
11	152	96.2	2332	2	US-08-670-707A-2
12	152	96.2	2332	3	US-09-037-601A-2
13	152	96.2	2332	3	US-09-324-867-3
14	152	96.2	2332	4	US-09-315-179-2
15	152	96.2	2332	4	US-09-523-656-2
16	152	96.2	2332	5	PCT-US93-03275-4
17	152	96.2	2332	5	PCT-US94-13200-2
18	152	96.2	2351	1	US-08-121-202-2
19	152	96.2	2351	1	US-08-366-851A-2
20	152	96.2	2351	6	5171844-2
21	152	96.2	2351	6	5422260-1
22	144	91.1	2304	3	US-09-324-867-4
23	144	91.1	2319	1	US-08-212-133A-8
24	144	91.1	2319	1	US-08-474-503-6
25	144	91.1	2319	2	US-08-670-707A-6
26	144	91.1	2319	3	US-09-037-601-6
27	144	91.1	2319	4	US-09-315-179-6

28 144 91.1 2319 4 US-09-523-656-28 Sequence 28, Appli
 29 144 91.1 2319 5 PCT-US94-13200-6 Sequence 6, Appli
 30 143 90.5 2343 3 US-09-324-867-2 Sequence 2, Appli
 31 136 86.1 541 1 US-08-121-202-4 Sequence 4, Appli
 32 136 86.1 1443 2 US-08-670-707A-39 Sequence 39, Appli
 33 136 86.1 1443 3 US-09-037-601-39 Sequence 39, Appli
 34 136 86.1 1443 4 US-09-315-179-39 Sequence 39, Appli
 35 136 86.1 1467 4 US-09-523-656-38 Sequence 38, Appli
 36 136 86.1 2115 3 US-09-324-867-5 Sequence 5, Appli
 37 136 86.1 2133 2 US-08-670-707A-37 Sequence 37, Appli
 38 136 86.1 2133 3 US-09-037-601-37 Sequence 37, Appli
 39 136 86.1 2133 4 US-09-315-179-37 Sequence 37, Appli
 40 136 86.1 2133 4 US-09-523-656-30 Sequence 30, Appli
 41 66 41.8 430 4 US-09-149-476-707 Sequence 707, Appli
 42 55 34.8 283 4 US-09-328-352-6473 Sequence 6473, Ap
 43 52 32.9 635 4 US-09-328-352-8029 Sequence 8029, Ap
 44 51 32.3 322 4 US-09-134-001C-2946 Sequence 2946, Ap
 45 48 30.4 186 1 US-07-960-981-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-209-916-1
 ; Sequence 1, Application US/09209916
 ; Patent No. 6358703
 ; GENERAL INFORMATION:
 ; APPLICANT: Cho, Myung-Sam
 ; APPLICANT: Chan, Sham-Yuen
 ; APPLICANT: Kelsey, William
 ; APPLICANT: Yee, Helena
 ; TITLE OF INVENTION: Expression System for Factor VIII
 ; FILE REFERENCE: MSB-7255
 ; CURRENT APPLICATION NUMBER: US/09/209,916
 ; CURRENT FILING DATE: 1998-12-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1438
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Derived from
 ; OTHER INFORMATION: human factor VIII sequence
 US-09-209-916-1

Query Match 96.2%; Score 152; DB 4; Length 1438;
 Best Local Similarity 96.7%; Pred. No. 7.9e-15;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TFLTAQTLLMDLQGLLSCHISSHQHDGME 30
 |||||
 Db 292 TFLTAQTLLMDLQGLLSCHISSHQHDGME 321

RESULT 2

US-08-683-839B-3
 ; Sequence 3, Application US/08683839B
 ; Patent No. 5744326
 ; GENERAL INFORMATION:
 ; APPLICANT: Ill, Charles . R. et al.
 ; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
 ; TITLE OF INVENTION: Regulatory Sequences to Increase Expression of
 ; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-839B-3

```

```

Query Match          96.2%; Score 152; DB 1; Length 1471;
Best Local Similarity 96.7%; Pred. No. 81e-15;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TELTAQTLLMDLGQFLLSCHSHQHDGME 30
   |||||
Db 311 TELTAQTLLMDLGQFLLSCHSHQHDGME 340

```

```

RESULT 3
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-882-083-2
; Query Match          96.2%; Score 152; DB 2; Length 1661;
; Best Local Similarity 96.7%; Pred. No. 9.4e-15;
; Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TELTAQTLLMDLGQFLLSCHSHQHDGME 30
   |||||
Db 311 TELTAQTLLMDLGQFLLSCHSHQHDGME 340

```

```

RESULT 4
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEPH: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-107-2

```

```

Query Match          96.2%; Score 152; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 9.4e-15;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TELTAQTLLMDLGQFLLSCHSHQHDGME 30
   |||||
Db 311 TELTAQTLLMDLGQFLLSCHSHQHDGME 340

```

```

RESULT 5
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-243-539-2

Query Match          96.2%; Score 152; DB 3; Length 1661;
Best Local Similarity 96.7%; Pred. No. 9.4e-15;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGFLLSCHISSHQHDGME 30
Db 311 TFLTAQTLLMDLGGFLLSCHISSHQHDGME 340

RESULT 6
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patricia L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4

Query Match          96.2%; Score 152; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGFLLSCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGGFLLSCHISSHQHDGME 321

RESULT 7
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
```

```

; TISSUE TYPE: Liver cDNA sequence
US-08-251-937A-4

Query Match          96.2%; Score 152; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGFLLSCHSHQHDGME 30
   |||||
Db 292 TFLTAQTLMDLQGFLLSCHSHQHDGME 321

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
US-08-212-133A-2

Query Match          96.2%; Score 152; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGFLLSCHSHQHDGME 30
   |||||
Db 292 TFLTAQTLMDLQGFLLSCHSHQHDGME 321

RESULT 9
US-08-276-594A-2

; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-594A-2

Query Match          96.2%; Score 152; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGFLLSCHSHQHDGME 30
   |||||
Db 292 TFLTAQTLMDLQGFLLSCHSHQHDGME 321

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:

```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
US-08-474-503-2

```

```

Query Match 96.2%; Score 152; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 321

```

```

RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids

```

```

; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-08-670-707A-2

```

```

Query Match 96.2%; Score 152; DB 2; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 321

```

```

RESULT 12
US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids

```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

Query Match 96.2%; Score 152; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 321

RESULT 13

US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632

GENERAL INFORMATION:

; APPLICANT: Lilliecrap, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 62516321ey, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.0010002/JAG/BJD
; CURRENT APPLICATION NUMBER: US/09/324,867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-059
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-324-867-3

Query Match 96.2%; Score 152; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
|||||
Db 293 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 322

RESULT 14

US-09-315-179-2

; Sequence 2, Application US/09315179
; Patent No. 6376463

GENERAL INFORMATION:

; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200

;
; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-315-179-2

Query Match 96.2%; Score 152; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 321

RESULT 15

US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563

GENERAL INFORMATION:

; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-523-656-2

Query Match 96.2%; Score 152; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.4e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHISSHQHDGME 321

Search completed: October 21, 2003, 18:59:48
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:53:43 ; Search time 21.25 Seconds
(without alignments)
236.415 Million cell updates/sec

Title: FVIII_PHE328S

Perfect score: 158

Sequence: 1 TFLTAQLLMDLGQFLLSCHSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	96.2	1438	14	US-10-006-091-1
2	152	96.2	1438	14	US-10-047-257-1
3	152	96.2	1438	15	US-10-225-900-1
4	152	96.2	1471	14	US-10-095-718-2
5	152	96.2	2332	10	US-09-957-641-2
6	152	96.2	2332	12	US-10-131-510A-2
7	152	96.2	2332	15	US-10-187-319-2
8	152	96.2	2351	12	US-10-133-907-4
9	152	96.2	2351	15	US-10-132-829-4
10	152	96.2	2351	15	US-10-172-712-27
11	144	91.1	2319	15	US-10-131-510A-6
12	144	91.1	2319	15	US-10-187-319-6
13	143	90.5	1431	14	US-10-095-718-4
14	136	86.1	1443	12	US-10-131-510A-39
15	136	86.1	1443	15	US-10-187-319-39

16	136	86.1	2133	12	US-10-131-510A-37
17	136	86.1	2133	15	US-10-187-319-37
18	75	47.5	1160	12	US-10-137-870-234
19	75	47.5	1160	12	US-10-140-018-234
20	75	47.5	1160	12	US-10-140-021-234
21	75	47.5	1160	12	US-10-140-274-234
22	75	47.5	1160	12	US-10-140-471-234
23	75	47.5	1160	12	US-10-140-807-234
24	75	47.5	1160	12	US-10-140-922-234
25	75	47.5	1160	12	US-10-140-926-234
26	75	47.5	1160	12	US-10-140-926-234
27	75	47.5	1160	12	US-10-141-698-234
28	75	47.5	1160	12	US-10-141-702-234
29	75	47.5	1160	12	US-10-141-704-234
30	75	47.5	1160	12	US-10-142-421-234
31	75	47.5	1160	12	US-10-142-432-234
32	75	47.5	1160	12	US-10-142-767-234
33	75	47.5	1160	12	US-10-143-033-234
34	75	47.5	1160	12	US-10-144-994-234
35	75	47.5	1160	12	US-10-145-628-234
36	75	47.5	1160	12	US-10-145-631-234
37	75	47.5	1160	12	US-10-145-633-234
38	75	47.5	1160	12	US-10-145-746-234
39	75	47.5	1160	12	US-10-145-748-234
40	75	47.5	1160	12	US-10-145-823-234
41	75	47.5	1160	12	US-10-145-826-234
42	75	47.5	1160	12	US-10-145-870-234
43	75	47.5	1160	12	US-10-145-876-234
44	75	47.5	1160	12	US-10-145-959-234
45	75	47.5	1160	12	US-10-146-724-234

ALIGNMENTS

RESULT 1
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-10-006-091-1

Query Match 96.2%; Score 152; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQLLMDLGQFLLSCHSHQHDGME 30
Db 292 TFLTAQLLMDLGQFLLSCHSHQHDGME 321

RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:

```

; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

```

```

Query Match          96.2%; Score 152; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
|||||

```

```

RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US2003007752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

```

```

Query Match          96.2%; Score 152; DB 15; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
|||||

```

```

RESULT 4
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony

```

```

; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

```

```

Query Match          96.2%; Score 152; DB 14; Length 1471;
Best Local Similarity 96.7%; Pred. No. 2.3e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||||
Db 311 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 340
|||||

```

```

RESULT 5
US-09-957-641-2
; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00
; CURRENT APPLICATION NUMBER: US/09/957,641
; CURRENT FILING DATE: 2001-09-16
; PRIOR APPLICATION NUMBER: US 60/234047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-957-641-2

```

```

Query Match          96.2%; Score 152; DB 10; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.7e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLLSCHSHQHDGME 321
|||||

```

```

RESULT 6
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601

```

```

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-510A-2

Query Match          96.2%; Score 152; DB 12; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.7e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLLSCHISSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLLSCHISSHQHDGME 321

RESULT 7
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: <Unknown>
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

Query Match          96.2%; Score 152; DB 15; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.7e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLLSCHISSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLLSCHISSHQHDGME 321

RESULT 8
US-10-133-907-4
; Sequence 4, Application US/10133907
; Publication No. US20030195223A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pall170
; CURRENT APPLICATION NUMBER: US/10/133,907
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-907-4

Query Match          96.2%; Score 152; DB 12; Length 2351;
Best Local Similarity 96.7%; Pred. No. 3.8e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLLSCHISSHQHDGME 30
Db 311 TELTAQTLLMDLGQFLLSCHISSHQHDGME 340

RESULT 9
US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pall170
; CURRENT APPLICATION NUMBER: US/10/132,829
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-4

```

Query Match 96.2%; Score 152; DB 15; Length 2351;
Best Local Similarity 96.7%; Pred. No. 3.8e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAOTLLMDLGQFLLSCHSHQHDGME 30
Db 311 TFLTAOTLLMDLGQFLLSCHSHQHDGME 340

RESULT 10

US-10-172-712-27
; Sequence 27, Application US/10172712
; Publication No. US20030125232A1
; GENERAL INFORMATION:
; APPLICANT: GRIFFIN, JOHN H.
; APPLICANT: GALE, ANDREW J.
; APPLICANT: GETZOFF, ELIZABETH D.
; APPLICANT: PELLEQUER, JEAN-LUC
; TITLE OF INVENTION: STABILIZED PROTEINS WITH ENGINEERED DISULFIDE BONDS
; FILE REFERENCE: 4198-4001U81
; CURRENT FILING DATE: 2002-09-30
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 60/298,578
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-172-712-27

Query Match 96.2%; Score 152; DB 15; Length 2351;
Best Local Similarity 96.7%; Pred. No. 3.8e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAOTLLMDLGQFLLSCHSHQHDGME 30
Db 311 TFLTAOTLLMDLGQFLLSCHSHQHDGME 340

RESULT 11

US-10-131-510A-6
; Sequence 6, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2319
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-131-510A-6

Query Match 91.1%; Score 144; DB 12; Length 2319;
Best Local Similarity 90.0%; Pred. No. 5.4e-12;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAOTLLMDLGQFLLSCHSHQHDGME 30
Db 312 TFLTAOTLLMDLGQFLLSCHSHQHDGME 341

RESULT 12

US-10-187-319-6
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; TITLE: Sequence of the Murine Factor VIII cDNA
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-187-319-6

Query Match 91.1%; Score 144; DB 15; Length 2319;
Best Local Similarity 90.0%; Pred. No. 5.4e-12;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAOTLLMDLGQFLLSCHSHQHDGME 30
Db 312 TFLTAOTLLMDLGQFLLSCHSHQHDGME 341

RESULT 13

US-10-095-718-4
; Sequence 4, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Bursstein, Haim

```

; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: canine B-domain deleted factor VIII
US-10-095-718-4

Query Match          90.5%; Score 143; DB 14; Length 1431;
Best Local Similarity 90.0%; Pred. No. 4.4e-12;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLSCHSHSHQHDGME 30
Db 306 TFLTAQTFLMDLQGLFLFCHIPSHQHDGME 335

RESULT 14
US-10-131-510A-39
; Sequence 39, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: factor VIII lacking
; OTHER INFORMATION: the B domain
US-10-131-510A-39

Query Match          86.1%; Score 136; DB 12; Length 1443;
Best Local Similarity 86.7%; Pred. No. 4.7e-11;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLSCHSHSHQHDGME 30
Db 312 TFLTAQTFLMDLQGLFLFCHSHSHHGME 341

```

```

RESULT 15
US-10-187-319-39
; Sequence 39, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-10-187-319-39

Query Match          86.1%; Score 136; DB 15; Length 1443;
Best Local Similarity 86.7%; Pred. No. 4.7e-11;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLSCHSHSHQHDGME 30
Db 312 TFLTAQTFLMDLQGLFLFCHSHSHHGME 341

```

Search completed: October 21, 2003, 19:01:20
Job time : 21.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:49:52 ; Search time 17.5 Seconds
(without alignments)
164.861 Million cell updates/sec

Title: FVIII_PHE328S

Perfect score: 158

Sequence: 1 TFLTACTLLMDLGGFLLSCHISSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	96.2	2351	1 EZHU	coagulation factor
2	144	91.1	2319	2 A47004	coagulation factor
3	136	86.1	2133	2 T42763	coagulation factor
4	62.5	39.6	1069	1 KUHU	ferroxidase (EC 1.1.1.1)
5	60	38.0	1059	1 A35210	ferroxidase (EC 1.1.1.1)
6	59	37.3	721	2 H82528	L-ascorbate oxidase
7	52.5	33.2	137	2 T07891	protein kinase (EC 2.7.1.1)
8	52	32.9	609	1 KSASL1	laccase (EC 1.10.3.1)
9	51	32.3	427	2 A12785	seryl-tRNA synthetase
10	51	32.3	435	2 C97565	seryl-tRNA synthetase
11	49	31.0	527	2 C70397	periplasmic cell d
12	48.5	30.7	273	2 C83968	hypothetical prote
13	48	30.4	379	2 F82963	hypothetical prote
14	48	30.4	398	2 S59055	myc2 protein - rai
15	48	30.4	406	2 A48059	oncoprotein zc-Myc
16	48	30.4	414	1 TVTRMC	transforming prote
17	48	30.4	438	2 JCL179	transforming prote
18	48	30.4	439	1 TVHOM	transforming prote
19	48	30.4	439	1 TVMS	transforming prote
20	48	30.4	439	1 TVTRMC	transforming prote
21	48	30.4	439	2 JCL178	transforming prote
22	48	30.4	439	2 JU0449	c-myc protein - ch
23	48	30.4	440	4 TVHUT	transforming prote
24	48	30.4	465	2 S03325	transforming prote
25	48	30.4	487	2 D71423	hypothetical prote
26	47.5	30.1	304	2 S76572	hypothetical prote
27	47	29.7	143	2 H64000	hypothetical prote
28	47	29.7	379	2 AH0542	conserved hypotet
29	47	29.7	519	2 H97724	multidrug resistan

ALIGNMENTS

RESULT 1

EZHU

coagulation factor VIII precursor [validated] - human
N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant c
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence-revision 28-Aug-1985 #text-change 08-Dec-2000
C:Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;
R:Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.
A:Reference number: I54318; MUID:93265012; PMID:1303178
A:Accession: I54318
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1921, 'S', 1923-2351 <RES>
A:Cross-references: GB:M88648; NID:G182381; PIDN:AAA52420.1; PID:G182393
R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; See
Nature 312, 330-337, 1984
A:Title: Expression of active human factor VIII from recombinant DNA clones.
A:Reference number: A00525; MUID:85061548; PMID:6438526
A:Accession: A00525
A:Molecule type: mRNA
A:Residues: 1-2351 <MOO>
A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.
S, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A:Reference number: I58059; MUID:85061550; PMID:6438528
A:Accession: I58059
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
A:Cross-references: GB:K01740; NID:G182802; PIDN:AAA52484.1; PID:G182803
R:Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.;
B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.
DNA 4, 333-349, 1985
A:Title: Characterization of the polypeptide composition of human factor VIII:C and
A:Reference number: A23584; MUID:86081164; PMID:3935400
A:Accession: A23584
A:Molecule type: mRNA
A:Residues: 1-2351 <TRU>
A:Cross-references: GB:M14113; NID:G182817; PIDN:AAA52485.1; PID:G182818
R:Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleave
ity.
A:Reference number: A26174; MUID:86159740; PMID:3082357
A:Accession: A26174
A:Molecule type: protein
A:Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <I
R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
L-ascorbate oxidas
subtilisin-like pr
indole-3-glycerol-
capsid precursor -
hypothetical prote
hypothetical prote
probable dehydroge
conserved hypotet
conserved hypotet
hypothetical prote
probable copper ox
hypothetical prote

Biochemistry 31, 3315-3325, 1992
 A;Title: Identification and functional importance of tyrosine sulfate residues within re
 A;Reference number: A42348; MUID:92207952; PMID:1554716
 A;Accession: A42348
 A;Molecule type: protein
 A;Residues: 20-36;336-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
 A;Experimental source: recombinant material from Chinese hamster ovary cells
 A;Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 R;Fay, P.J.; Smudzin, T.M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A;Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A;Reference number: A43986; MUID:89340500; PMID:2503509
 A;Accession: A43986
 A;Molecule type: protein
 A;Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>
 R;Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
 J. Biol. Chem. 266, 740-746, 1991
 A;Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A;Reference number: A56109; MUID:91093266; PMID:1898735
 A;Contents: annotation; sulfation
 R;Gitschier, J.; Wood, W.I.; Goralika, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
 Nature 312, 326-330, 1984
 A;Title: Characterization of the human factor VIII gene.
 A;Reference number: A56196; MUID:85061547; PMID:6438525
 A;Contents: annotation; introns
 R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A;Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
 A;Reference number: A56216; MUID:95338127; PMID:7613471
 A;Contents: annotation; disulfide bonds
 A;Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
 R;Kjallke, M.; Hedner, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A;Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A;Reference number: S63527; MUID:96163459; PMID:8575434
 A;Accession: S63527
 A;Molecule type: protein
 A;Residues: 733-752;753-759 <KJA>
 R;Lind, P.; Larsson, K.; Spira, J.; Sydow-Baekman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A;Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A;Reference number: S66445; MUID:96048024; PMID:7556150
 A;Accession: S66445
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1668-1685 <LIN>
 C;Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure p
 C;Genetics:
 A;Gene: GDB:F8C
 A;Cross-references: GDB:119124; OMIM:306700
 A;Map position: Xg28-Xq28
 A;Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C;Function:
 A;Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 A;Pathway: blood coagulation
 C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
 F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>
 F:20-356/Domain: A1 <DA1>
 F:23-348/Domain: ferroxidase repeat homology <FO1>
 F:392-759/Domain: A2 <DA2>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:760-1667/Domain: B <DBO>
 F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>
 F:1709-2038/Domain: A3 <DA3>
 F:1716-2038/Domain: ferroxidase repeat homology <FO3>
 F:2039-2191/Domain: C1 <DC1>
 F:2039-2188/Domain: discoidin I amino-terminal homology <DN1>
 F:2192-2351/Domain: C2 <DC2>
 F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>
 F:60,258,601,776,803,847,919,962,982,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #st
 F:355-356/Cleavage site: Arg-Wet (coagulation factor Xa, protein C) #status prediacte
 F:365-737/738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimen
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen
 F:414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimen
 F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experim
 F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
 F:2193-2345/Disulfide bonds: #status predicted
 Query Match 96.2%; Score 152; DB 1; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 4.9e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLQGFLLSHSHQHDGME 30
 Db 311 TELTAQTLLMDLQGFLLSHSHQHDGME 340
 RESULT 2
 A47004
 coagulation factor VIII precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 09-Sep-1994 #sequence-revision 09-Sep-1994 #text_change 18-Jun-1999
 C;Accession: A47004
 R;Elder, B.; Lakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993
 A;Title: Sequence of the murine factor VIII cDNA.
 A;Reference number: A47004; MUID:93300511; PMID:8314577
 A;Accession: A47004
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-2319 <ELD>
 A;Cross-references: GB:L05573; NID:g192456; PIDN:AAA37385.1; PID:g192457
 C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:23-349/Domain: ferroxidase repeat homology <FO1>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:1686-2006/Domain: ferroxidase repeat homology <FO3>
 F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>
 F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>
 Query Match 91.1%; Score 144; DB 2; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 7.4e-13;
 Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLQGFLLSHSHQHDGME 30
 Db 312 TELTAQTLLMDLQGFLLSHSHQHDGME 341
 RESULT 3
 T42763
 coagulation factor VIII precursor - pig
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 11-Jan-2000 #sequence-revision 11-Jan-2000 #text_change 09-Jun-2000
 C;Accession: T42763
 R;Lollar, P.
 submitted to the EMBL Data Library, August 1996
 A;Reference number: 422269
 A;Accession: T42763
 A;Status: preliminary; translated from GB/EMBL/DBDJB
 A;Molecule type: mRNA
 A;Residues: 1-2133 <LOL>
 A;Cross-references: EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AAB06705.1
 C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi
 C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A;
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
 F:23-349/Domain: ferroxidase repeat homology <FO1>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

100

A:Map position: circular chromosome
C:Superfamily: serine-tRNA ligase

Query Match 32.3%; Score 51; DB 2; Length 427;
Best Local Similarity 42.3%; Pred. No. 9.3;
Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 1 TFLTAQTLLMD--LGQFLLSCHISSH 24
| | | | : : | | | | : : | |
Db 165 TVLTSQLARLERALQGFMDLHTSEH 190

RESULT 10

C97565
seryl-tRNA synthetase (serine-tRNA ligase) (serres) [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: C97565

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: C97565

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-435 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87476.1; PID:gl15156798; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C3129

A:Map position: circular chromosome

C:Superfamily: serine-tRNA ligase

Query Match 32.3%; Score 51; DB 2; Length 435;
Best Local Similarity 42.3%; Pred. No. 9.5;
Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 1 TFLTAQTLLMD--LGQFLLSCHISSH 24
| | | | : : | | | | : : | |
Db 173 TVLTSQLARLERALQGFMDLHTSEH 198

RESULT 11

C70397
periplasmic cell division protein (SufI) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70397

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: C70397

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-527 <AQF>

A:Cross-references: GB:AE000724; NID:Q2983585; PIDN:AAC07157.1; PID:g2983586; GB:AE00065

A:Experimental source: strain VF5

C:Genetics:

A:Gene: sufi

Query Match 31.0%; Score 49; DB 2; Length 527;
Best Local Similarity 53.3%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 15 FLLSCHISSHQHDM 29
: | | | | | : | | |
Db 505 YLLCHLHLEHDM 519

RESULT 12

GB83968

hypothetical protein BH2551 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83968
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83968

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-273 <STO>

A:Cross-references: GB:AP001515; GB:BA000004; NID:gl0174886; PIDN:BAB06270.1; GSPDB

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2551

C:Superfamily: Streptomyces coelicolor hypothetical protein SC4A10.14c

Query Match 30.7%; Score 48.5; DB 2; Length 273;

Best Local Similarity 38.2%; Pred. No. 14;

Matches 13; Conservative 5; Mismatches 5; Indels 11; Gaps 2;

QY 6 QTLMDLG---QFLLSCHIS-----SHQHDG 28
: | | | : | | | : | | | : | | | : | | | : | | |

Db 226 EOLLIDAGLEKAQLVSSHCSAQSNLLFSHRHEG 259
: | | | : | | | : | | | : | | | : | | | : | | |

RESULT 13

F82963

hypothetical protein PA5471 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: F82963

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Miziochichi, S.D.; Warriner, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; ; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic

A:Reference number: A82950; MUID:2043737; PMID:10984043

A:Accession: F82963

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <STO>

A:Cross-references: GB:AE004959; GB:AE004091; NID:g9951791; PIDN:AAG08856.1; GSPDB:

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA5471

C:Superfamily: Streptomyces coelicolor hypothetical protein SCE15.16c

Query Match 30.4%; Score 48; DB 2; Length 379;

Best Local Similarity 50.0%; Pred. No. 23;

Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 12 LGQFLLSCHISSHQHDM 29
| | | : | | | : | | | : | | |

Db 174 LGQATLQEHVSLHGAGL 191

RESULT 14

S59905

myc2 protein - rainbow trout

C:Species: Oncorhynchus mykiss (rainbow trout)

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 16-Jul-1999

C:Accession: S59905

R:Panno, J.P.; McKeown, B.A.

Biochim. Biophys. Acta 1264, 7-11, 1995

A:Title: Cloning and expression of a myc family member from the pituitary gland of

A:Reference number: S59905; MUID:96038810; PMID:7578259

A:Accession: S59905

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-398 <PAN>

A:Cross-references: GB:S79770; NID:gl176453; PIDN:AAB35487.1; PID:gl176454

A;Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 2-Leu,
C;Genetics:

A;Gene: myc2
C;Superfamily: myc transforming protein; myc transforming protein homology
C;Keywords: proto-oncogene
F;12-394/Domain: myc transforming protein homology <MYC>

Query Match 30.4%; Score 48; DB 2; Length 398;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 LLSCHISSHQHD 27
| | | | |
Db 247 LKRCVSTHQHN 258

RESULT 15

A48059

oncoprotein zc-Myc - zebra fish

C;Species: Brachydanio rerio (zebra fish)

C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999

C;Accession: A48059

R;Schreiber-Agus, N.; Horner, J.; Torres, R.; Chiu, P.C.; DePinho, R.A.

Mol. Cell. Biol. 13, 2765-2775, 1993

A;Title: Zebra fish myc family and max genes: differential expression and oncogenic activation

A;Reference number: A48059; MUID:93233639; PMID:847440

A;Accession: A48059

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA; DNA

A;Residues: 1-406 <SCH>

A;Cross-references: GB:L11710; NID:g215041; PIDN:AAA02482.1; FID:g215042

A;Note: sequence extracted from NCBI backbone (NCBIP:129810)

C;Superfamily: myc transforming protein; myc transforming protein homology

F;14-405/Domain: myc transforming protein homology <MYC>

Query Match 30.4%; Score 48; DB 2; Length 406;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 LLSCHISSHQHD 27
| | | | |
Db 257 LKRCVSTHQHN 268

Search completed: October 21, 2003, 18:58:38
Job time : 18.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:22 : Search time 9.25 Seconds
(without alignments)
152.519 Million cell updates/sec

Title: FVIII_PHE328S

Perfect score: 158

Sequence: 1 TFLTAQTLLMDLGGFLHSCHTSSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	152	96.2	2351	1	FA8_HUMAN
2	144	91.1	2319	1	FA8_MOUSE
3	136	86.1	2133	1	FA8_PIG
4	62.5	39.6	1065	1	CERU_HUMAN
5	61	38.6	1062	1	CERU_MOUSE
6	60	38.0	1059	1	CERU_RAT
7	52	32.9	609	1	LAC1_EMENI
8	51	32.3	427	1	SVS_AGRUS
9	49.5	31.3	281	1	COX3_RHIST
10	48	30.4	93	1	FXI8_HUMAN
11	48	30.4	394	1	MYC1_CYPFA
12	48	30.4	399	1	MYC_CARAU
13	48	30.4	401	1	MYC2_CYPFA
14	48	30.4	406	1	MYC_BRABE
15	48	30.4	414	1	MYC_ONCMY
16	48	30.4	438	1	MYC_CALGA
17	48	30.4	439	1	MYC_CANFA
18	48	30.4	439	1	MYC_HUMAN
19	48	30.4	439	1	MYC_HYLLA
20	48	30.4	439	1	MYC_WARMO
21	48	30.4	439	1	MYC_MOUSE
22	48	30.4	439	1	MYC_PANTR
23	48	30.4	439	1	MYC_PIG
24	48	30.4	439	1	MYC_RAT
25	48	30.4	439	1	MYC_SHEEP
26	48	30.4	1134	1	IF3X_HUMAN
27	47.5	30.1	304	1	IY99_SYNY3
28	47	29.7	143	1	YIFN_HAEN
29	47	29.7	642	1	PRSA_STRAT
30	46	29.1	486	1	LAC1_BOTCI
31	45.5	28.8	266	1	TRPC_METJA
32	45.5	28.8	797	1	T4AQ_HUMAN
33	45	28.5	437	1	CTBQ_XENLA

34	45	28.5	552	1	ASO_CUCPM	P37064	cucurbita p
35	45	28.5	578	1	ASO_TOBAC	Q40588	nicotiana t
36	45	28.5	579	1	ASO_CUCMA	P24792	cucurbita m
37	45	28.5	587	1	ASO_CUCSA	P14133	cucumis sat
38	45	28.5	608	1	YD56_YEAST	Q04399	saccharomyc
39	45	28.5	622	1	YAK8_SCHPO	Q09920	schizosacch
40	45	28.5	839	1	V2A_CMVQ	P06012	arabacia lix
41	44.5	28.2	109	1	COX3_ARBLI	Q33752	arabacia lix
42	44.5	28.2	205	1	FLRE_HUMAN	P30043	homo sapien
43	44.5	28.2	213	1	ATT2_VACCV	P24758	vaccinia vi
44	44.5	28.2	261	1	COX3_PARLI	P12702	paracentrot
45	44.5	28.2	322	1	ATT2_VACCC	P21114	vaccinia vi

ALIGNMENTS

RESULT 1
FA8_HUMAN
AC P00451; STANDARD; PRT: 2351 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component)
DE (Antithrombin factor) (AHF).
GN F8 OR F8C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Maslarz F.R., Merryweather J.P., Najarian R.,
RA Pacht C., Porter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA Nordfang O., Ezban M.;
RT "Characterization of the polypeptide composition of human factor
VIII:C and the nucleotide sequence and expression of the human kidney
cDNA.";
RL DNA 4:333-349(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wilson K.L.,
RA Delwart E., Tuddenham E.G.D., Vohar G.A., Lawn R.M.;
RT "Expression of active human factor VIII from recombinant DNA clones.";
RL Nature 312:330-337(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Soltzman L.A., Buecker J.L.,
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA Hewick R.M.;
RT "Molecular cloning of a cDNA encoding human antihaemophilic factor.";
RL Nature 312:342-347(1984).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93265012; PubMed=1303178;
RA Gitschier J., Wood W.I.;
RT "Sequence of the exon-containing regions of the human factor VIII
gene.";
RL Hum. Mol. Genet. 1:199-200(1992).
RN [5]
RP SEQUENCE OF 2064-2070 FROM N.A.
RA de Water N.S., Williams R., Browett P.J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SULFATION OF TYR-1699.

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor.";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP VARIANT HEMA GLY-291.
 RX MEDLINE=92207952; PubMed=1554716;
 RA Pittman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT HEMA GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac.";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT HEMA PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT HEMA GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT HEMA GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu->Gly substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT HEMA LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT HEMA LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HEMA HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT HEMA CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANT HEMA GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT HEMA CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANT HEMA PHE-1699 AND CYS-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:65-71(1990).
 RN [25]
 RP VARIANT HEMA CYS-1728 AND ASP-1941.
 RX MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";

Query Match 96.2%; Score 152; DB 1; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 5.2e-14;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLLSCHISSHQDGM 30
 |||||:|||||:|||||:|||||:|||||:|||||
 Db 311 TFLTAQTLLMDLQGLFLLSCHISSHQDGM 340

RESULT 2

FA8_MOUSE STANDARD; PRT; 2319 AA.

AC Q06194;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Coagulation factor VIII precursor (Procoagulant component).
 GN F8 OR CF8 OR F8C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CHA; TISSUE=Liver;
 RX MEDLINE=93300511; PubMed=8314577;
 RA Elder B., Lakich D., Gitschler J.;

RT "Sequence of the murine factor VIII cDNA";
 RL Genomics 16:374-379(1993).
 CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR XA.

CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L05573; AAA37385.1; -.
 DR PIR; A47004; A47004.

DR HSP; P00451; 1CRG.

DR MGD; MGI:88383; F8.

DR InterPro; IPR001117; Cu-oxidase.

DR InterPro; IPR000421; FA58_C.

DR Pfam; PF00394; Cu-oxidase; 3.

DR Pfam; PF00754; F5_F8_type_C; 2.

DR SMART; SM00231; FA58C; 2.

DR PROSITE; PS01285; FA58C_1; 2.

DR PROSITE; PS01286; FA58C_2; 2.

DR PROSITE; PS00022; FA58C_3; 2.

DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.

KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;

KW Signal; Glycoprotein; Sulfation.

FT SIGNAL 1 19

FT CHAIN 20 2319

FT DOMAIN 20 349

FT DOMAIN 20 199

FT DOMAIN 207 349

FT DOMAIN 207 349

FT DOMAIN 399 730

FT DOMAIN 399 730

FT DOMAIN 583 730

FT DOMAIN 760 1640

FT DOMAIN 1683 2008

FT DOMAIN 1683 1845

FT DOMAIN 1855 2008
 FT DOMAIN 2008 2156
 FT DOMAIN 2156 2313
 FT SITE 391 392
 FT SITE 759 760
 FT SITE 1678 1679
 FT SITE 1324 1325
 FT SITE 1640 1641
 FT MOD_RES 367 367
 FT MOD_RES 737 737
 FT MOD_RES 738 738
 FT MOD_RES 742 742
 FT MOD_RES 1669 1669
 FT MOD_RES 1687 1687
 FT DISULFID 173 199
 FT DISULFID 547 573
 FT DISULFID 1819 1845
 FT DISULFID 2008 2156
 FT DISULFID 2161 2313
 FT CARBOHYD 61 61
 FT CARBOHYD 233 233
 FT CARBOHYD 259 259
 FT CARBOHYD 423 423
 FT CARBOHYD 601 601
 FT CARBOHYD 880 880
 FT CARBOHYD 958 958
 FT CARBOHYD 1015 1015
 FT CARBOHYD 1022 1022
 FT CARBOHYD 1026 1026
 FT CARBOHYD 1044 1044
 FT CARBOHYD 1076 1076
 FT CARBOHYD 1087 1087
 FT CARBOHYD 1136 1136
 FT CARBOHYD 1161 1161
 FT CARBOHYD 1192 1192
 FT CARBOHYD 1255 1255
 FT CARBOHYD 1268 1268
 FT CARBOHYD 1273 1273
 FT CARBOHYD 1274 1274
 FT CARBOHYD 1302 1302
 FT CARBOHYD 1316 1316
 FT CARBOHYD 1340 1340
 FT CARBOHYD 1378 1378
 FT CARBOHYD 1797 1797
 FT CARBOHYD 2105 2105
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;

Query Match

ID FA8_PIG 91.1%; Score 144; DB 1; Length 2319;

Best Local Similarity 90.0%; Pred. No. 7.5e-13;

Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLFLLSCHISSHQDGM 30

|||||:|||||:|||||:|||||:|||||:|||||

Db 312 TFLTAQTLLMDLQGLFLLSCHISSHQDGM 341

RESULT 3

FA8_PIG

AC P12263; Q95243; STANDARD; PRT; 2133 AA.

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Coagulation factor VIII precursor (Procoagulant component).

GN F8 OR CF8.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RA Healey J.F., Lubin I.M., Lollar P.;

Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

[2]
 SEQUENCE OF 705-1573 FROM N.A.
 MEDLINE=86287369; PubMed=3016730;
 Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
 Kaufman R.J.;
 "A large region (approximately equal to 95 kDa) of human factor VIII
 is dispensable for in vitro procoagulant activity.";
 Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
 [3]
 SEQUENCE OF 392-759 FROM N.A.
 MEDLINE=94179260; PubMed=7510693;
 Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
 "Elimination of a major inhibitor epitope in factor VIII.";
 J. Biol. Chem. 269:8639-8641(1994).
 CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS PROTHROMBIN TO THE
 ACTIVATED FORM, FACTOR XA.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U49517; AAB06705.1; -
 DR PIR: A25945; A25945.
 DR PIR: T42763; T42763.
 DR HSP: P00451; LCFG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS01285; FA58C.1; 2.
 DR PROSITE: PS01286; FA58C.2; 2.
 DR PROSITE: PS00022; FA58C.3; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 KW Signal; Glycoprotein; Sulfation.
 FT SIGNAL 1 19
 FT CHAIN 20 2133
 FT DOMAIN 20 357
 FT F5/8 TYPE A 1.
 FT DOMAIN 20 199
 FT PLASTOCYANIN-LIKE 1.
 FT DOMAIN 207 357
 FT F5/8 TYPE A 2.
 FT DOMAIN 399 730
 FT PLASTOCYANIN-LIKE 3.
 FT DOMAIN 583 730
 FT PLASTOCYANIN-LIKE 4.
 FT DOMAIN 760 1599
 FT F5/8 TYPE A 3.
 FT DOMAIN 1495 1822
 FT PLASTOCYANIN-LIKE 5.
 FT DOMAIN 1495 1659
 FT PLASTOCYANIN-LIKE 6.
 FT DOMAIN 1669 1822
 FT F5/8 TYPE C 1.
 FT DOMAIN 1822 1970
 FT F5/8 TYPE C 2.
 FT SITE 391 392
 FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 759 760
 FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1449 1450
 FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1490 1491
 FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT MOD_RES 738 737
 FT CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT MOD_RES 742 742
 FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 173 199
 FT SULFATION (BY SIMILARITY).
 FT DISULFID 547 573
 FT SULFATION (BY SIMILARITY).
 FT DISULFID 1633 1659
 FT PROBABLE.
 FT DISULFID 1822 1970
 FT BY SIMILARITY.
 FT DISULFID 1975 2127
 FT BY SIMILARITY.
 FT CARBOHYD 233 233
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 86.1%; Score 136; DB 1; Length 2133;
 Best Local Similarity 86.7%; Pred. No. le-11;
 Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQLLSCHISSHHQDGM 30
 DB 312 TELTAQTLLMDLGGQLLSCHISSHHQDGM 341

RESULT 4
 CERU_HUMAN
 ID CERU_HUMAN STANDARD; PRT; 1065 AA.
 AC P00450; Q14063;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ceruloplasmin precursor (EC 1.16.3.1) (Ferroxidase).
 GN CP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86259737; PubMed=2873574;
 RA Koshchinsky M.I., Funk W.D., van Oost B.A., McGillivray R.T.A.;
 RT "Complete cDNA sequence of human ceruloplasmin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5086-5090(1986).
 RN [2]
 RP SEQUENCE OF 1-1006 FROM N.A.
 RX MEDLINE=95217183; PubMed=7702601;
 RA Daimon M., Yamatani K., Igarashi M., Fukase N., Kawanami T.,
 RA Kato T., Tomimaga M., Sasaki H.;
 RT "Fine structure of human ceruloplasmin gene.";
 RL Biochem. Biophys. Res. Commun. 208:1028-1035(1995).
 RN [3]
 RP SEQUENCE OF 1-40; 549-599; 784-829 AND 919-952 FROM N.A.
 RX MEDLINE=86275241; PubMed=3755405;
 RA Mercer J.F.B., Grimes A.;
 RT "Isolation of a human ceruloplasmin cDNA clone that includes the N-
 terminal leader sequence.";
 RL FEBS Lett. 203:185-190(1986).
 RN [4]
 RP SEQUENCE OF 218-1065 FROM N.A.
 RX MEDLINE=86205876; PubMed=3486416;
 RA Yang F., Naylor S.L., Lum J.B., Cutshaw S., McCombs J.L.,
 RA Naberhaus K.H., McGill J.R., Adrian G.S., Moore C.M., Barnett D.R.,

RA Bowman B.H.;
 RT "Characterization, mapping, and expression of the human ceruloplasmin
 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3257-3261(1986).
 RN [5]
 RP SEQUENCE OF 20-1065.
 RX MEDLINE=84119493; PubMed=6582496;
 RA Takahashi N., Orel T.L., Putnam F.W.;
 RT "Single-chain structure of human ceruloplasmin: the complete amino
 acid sequence of the whole molecule.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:390-394(1984).
 RN [6]
 RP SEQUENCE OF 158-333; 518-724 AND 858-1065.
 RX MEDLINE=83117800; PubMed=6571985;
 RA Takahashi N., Bauman R.A., Orel T.L., Dwulet F.E., Wang C.-C.,
 Putnam F.W.;
 RT "Internal triplication in the structure of human ceruloplasmin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:115-119(1983).
 RN [7]
 RP SEQUENCE OF 501-905.
 RX MEDLINE=81199407; PubMed=6940148;
 RA Dwulet F.E., Putnam F.W.;
 RT "Complete amino acid sequence of a 50,000-dalton fragment of human
 ceruloplasmin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:790-794(1981).
 RN [8]
 RP SEQUENCE OF 907-1065.
 RX MEDLINE=80137543; PubMed=6987229;
 RA Kingston I.B., Kingston B.L., Putnam F.W.;
 RT "Primary structure of a histidine-rich proteolytic fragment of human
 ceruloplasmin. I. Amino acid sequence of the cyanogen bromide
 peptides.";
 RL J. Biol. Chem. 255:2878-2885(1980).
 RN [9]
 RP SEQUENCE OF 907-1065.
 RX MEDLINE=80137544; PubMed=6987230;
 RA Kingston I.B., Kingston B.L., Putnam F.W.;
 RT "Primary structure of a histidine-rich proteolytic fragment of human
 ceruloplasmin. II. Amino acid sequence of the tryptic peptides.";
 RL J. Biol. Chem. 255:2886-2896(1980).
 RN [10]
 RP SEQUENCE OF 1007-1061 FROM N.A.
 RX MEDLINE=90285218; PubMed=2355023;
 RA Yang F.M., Friedrichs W.E., Cupples R.L., Banifacio M.J.,
 Sanford J.A., Horton W.A., Bowman B.H.;
 RT "Human ceruloplasmin. Tissue-specific expression of transcripts
 produced by alternative splicing.";
 RL J. Biol. Chem. 265:10780-10785(1990).
 RN [11]
 RP REVIEW
 RX MEDLINE=22049919; PubMed=12055353;
 RA Hellman N.E., Gitlin J.D.;
 RT "Ceruloplasmin metabolism and function.";
 RL Annu. Rev. Nutr. 22:439-458(2002).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
 RA Zaitseva I., Zaitsev V., Card G., Moshkov K., Bax B., Ralph A.,
 Lindley P.;
 RT "The x-ray structure of human serum ceruloplasmin at 3.1 A: nature of
 the copper centres.";
 RL J. Biol. Inorg. Chem. 1:15-23(1996).
 CC -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
 MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
 ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
 AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: 4 Fe(2+) + 4 H(+) + O(2) -> 4 Fe(3+) + 2 H(2)O.
 CC -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
 THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
 KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
 BINUCLEAR.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE
 PLASMA.
 CC -!- DISEASE: Defects in CP are the cause of aceruloplasminemia

CC [MM:604290], an autosomal recessive disorder of iron metabolism.
 CC It is characterized by iron accumulation in the brain as well as
 CC visceral organs. Clinical features consist of the triad of retinal
 CC degeneration, diabetes mellitus and neurological disturbances.
 CC -!- DISEASE: Ceruloplasmin levels are decreased in Wilson's disease,
 CC in which copper cannot be incorporated into ceruloplasmin in
 CC liver because of defects in the copper-transporting ATPase 2.
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M13699; AAA51976.1; -;
 CC EMBL; D45045; BAA08085.1; -;
 CC EMBL; D45044; BAA08084.1; -;
 CC EMBL; D45028; BAA08084.1; JOINED.
 CC EMBL; D45029; BAA08084.1; JOINED.
 CC EMBL; D45030; BAA08084.1; JOINED.
 CC EMBL; D45031; BAA08084.1; JOINED.
 CC EMBL; D45032; BAA08084.1; JOINED.
 CC EMBL; D45033; BAA08084.1; JOINED.
 CC EMBL; D45034; BAA08084.1; JOINED.
 CC EMBL; D45035; BAA08084.1; JOINED.
 CC EMBL; D45036; BAA08084.1; JOINED.
 CC EMBL; D45037; BAA08084.1; JOINED.
 CC EMBL; D45038; BAA08084.1; JOINED.
 CC EMBL; D45039; BAA08084.1; JOINED.
 CC EMBL; D45040; BAA08084.1; JOINED.
 CC EMBL; D45041; BAA08084.1; JOINED.
 CC EMBL; D45042; BAA08084.1; JOINED.
 CC EMBL; D45043; BAA08084.1; JOINED.
 CC EMBL; D00025; BAA00019.1; -;
 CC EMBL; X04135; CAA27752.1; -;
 CC EMBL; X04136; CAA27753.1; -;
 CC EMBL; X04137; CAA27754.1; -;
 CC EMBL; X04138; CAA27755.1; -;
 CC EMBL; M03536; AAA51975.1; -;
 CC EMBL; J05506; -; NOT_ANNOTATED_CDS.
 CC PIR; A25443; K0HU.
 CC PDB; 1KCW; 12-FEB-97.
 CC GlycositedB; P00450; -;
 CC SWISS-2DPAGE; P00450; HUMAN.
 CC Siena-2DPAGE; P00450; -;
 CC Genew; HGNC:2295; CP.
 CC MIM; 117700; -;
 CC MIM; 604290; -;
 CC GO; GO:0005615; C:extracellular space; TAS.
 CC GO; GO:0006879; P:iron ion homeostasis; TAS.
 CC InterPro; IPR001117; Cu-oxidase.
 CC InterPro; IPR002355; MultiCu_oxidase2.
 CC Pfam; PF00394; Cu-oxidase; 3.
 CC PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 CC PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 CC Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
 KW Signal; Polymorphism; 3D-structure.
 KW SIGNAL 1 19
 FT CHAIN 20 1065 CERULOPLASMIN.
 FT DOMAIN 20 357 F5/8 TYPE A 1.
 FT DOMAIN 20 200 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 209 357 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 370 718 F5/8 TYPE A 2.
 FT DOMAIN 370 560 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 570 718 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 730 1061 F5/8 TYPE A 3.
 FT DOMAIN 730 900 PLASTOCYANIN-LIKE 5.
 FT DOMAIN 908 1061 PLASTOCYANIN-LIKE 6.
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .).


```

FT SIGNAL          1    20    POTENTIAL.
FT CHAIN          21    609    LACCASE.
FT DOMAIN         45    141    PLASTOCYANIN-LIKE 1.
FT DOMAIN         270   372    PLASTOCYANIN-LIKE 2.
FT DOMAIN         463   602    PLASTOCYANIN-LIKE 3.
FT METAL          79    79    COPPER (TYPE 2) (PROBABLE).
FT METAL          81    81    COPPER (TYPE 3) (PROBABLE).
FT METAL          123   123    COPPER (TYPE 3) (PROBABLE).
FT METAL          125   125    COPPER (TYPE 3) (PROBABLE).
FT METAL          508   508    COPPER (TYPE 1) (PROBABLE).
FT METAL          511   511    COPPER (TYPE 2) (PROBABLE).
FT METAL          513   513    COPPER (TYPE 3) (PROBABLE).
FT METAL          585   585    COPPER (TYPE 3) (PROBABLE).
FT METAL          586   586    COPPER (TYPE 1) (PROBABLE).
FT METAL          587   587    COPPER (TYPE 3) (PROBABLE).
FT METAL          591   591    COPPER (TYPE 1) (PROBABLE).
FT METAL          596   596    COPPER (TYPE 1) (PROBABLE).
FT CARBOHYD       75    75    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       257   257    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       268   268    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       403   403    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       443   443    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       486   486    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       531   531    N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD       546   546    N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE       609 AA; 68037 MW; E667A42EF909512 CRC64;

Query Match          32.9%; Score 52; DB 1; Length 609;
Best Local Similarity 64.3%; Pred. No. 5.2;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 16 LLSCHISSHODGM 29
   :|||:|||||
Db 583 ILHCHIASHQMGGM 596

RESULT 8
SYS_AGRT5          STANDARD; PRT; 427 AA.
AC O8UFO2;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE SERYL-tRNA synthetase [EC 6.1.1.11] (Serine--tRNA ligase) (SerRS).
GN SERS OR ATU1703 OR AGR_C_3129.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse S., Saenphimmachak C., Wu Z., Romero P.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,

```

```

RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
CC + L-seryl-tRNA(Ser).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC FMBL; A8009126; AAL42703.1;
CC EMBL; A8008091; AAK87476.1; ALT_INIT.
CC PIR; C97565; C97565.
CC HAMAP; MF_00176; ; 1.
CC InterPro; IPR002314; tRNA-synt_2b.
CC InterPro; IPR002317; tRNA-synt_ser.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02403; Seryl_tRNA_N; 1.
CC Pfam; PF00587; tRNA-synt_2b; 1.
CC PRINTS; PR00981; TRNASYNTHSER.
CC TIGRPFMS; TIGR00414; sers; 1.
CC PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 427 AA; 47599 MW; 95FA03327A365C26 CRC64;

Query Match          32.3%; Score 51; DB 1; Length 427;
Best Local Similarity 42.3%; Pred. No. 5.2;
Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 1 TFLTAQTLLMD--LGGFLLSCHISSH 24
   |||:|:|:|:|:|:|:|
Db 165 TVLTSLARLERALGQFMIDLHTEH 190

RESULT 9
COX3_RHIST
ID COX3 RHIST      STANDARD; PRT; 281 AA.
AC P80441;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide III (EC 1.9.3.1).
GN COX3.
OS Rhizopus stolonifer (Rhizopus nigricans).
OC Mitochondrion.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
OX NCBI_TaxID=4846;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DAOM 14828; PubMed=7490780;
RX MEDLINE=96081490;
RA Paquin B., Forget L., Roewer I., Lang B.F.;
RT "Molecular phylogeny of Allomyces macrogynus: congruency between
RL nuclear ribosomal RNA- and mitochondrial protein-based trees."
RL J. Mol. Evol. 41:657-665(1995).
CC -!- FUNCTION: Subunits I, II and III form the functional core of
CC the enzyme complex.
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 3 family.
CC InterPro: IPR000298; CytC_oxidase_III.
DR Pfam; PF00510; COX3; 1.
DR ProDom; PD000362; CytC_oxidase_III; 1.

```


CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Carassius.
 CC NCBI_TaxID=7957;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC Zhang H., Zhang H., Okamoto N., Ikeda Y.;
 CC Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
 CC BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
 CC RECOGNIZES THE CORE SEQUENCE CAC[GA]TG. SEEMS TO ACTIVATE THE
 CC TRANSCRIPTION OF GROWTH-RELATED GENES.
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D31729; BAA06527.1; -
 CC HSSP; P01106; IA93.
 CC TRANSFAC; T03454; -
 CC InterPro; IPR001092; HLH_basic.
 CC InterPro; IPR003327; Myc-LZ.
 CC InterPro; IPR002418; TF_Myc.
 CC Pfam; PF00010; HLH; 1.
 CC Pfam; PF02344; Myc-LZ; 1.
 CC Pfam; PF01056; Myc-N-term; 1.
 CC PRINTS; PR00044; LEUZIPPRMYC.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HLH_1; FALSE_NEG.
 CC PROSITE; PS00888; HLH_2; 1.
 CC Nuclear protein; DNA-binding; Transcription regulation; Activator;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 198 206 POLY-SER.
 FT DOMAIN 207 231 POLY-GLU.
 FT DOMAIN 315 328 BASIC DOMAIN.
 FT DOMAIN 329 368 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 374 395 LEUCINE-ZIPPER (POTENTIAL).
 FT CARBOHYD 58 58 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 399 AA; 45756 MW; FC4B2CEC361F1D42 CRC64;
 Query Match 30.4%; Score 48; DB 1; Length 399;
 Best Local Similarity 58.3%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 16 LLSCHISSHQHD 27
 Db 259 LKRCHVSTHQHN 270
 | |||:||||
 RESULT 13
 MYC2_CYPCA
 ID MYC2_CYPCA STANDARD; PRT; 401 AA.
 AC Q90342;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE MYC II protein (c-myc II).
 GN CAM2.
 OS Cyprinus carpio (Common carp).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Cyprinus.
 CC NCBI_TaxID=7962;
 CC [1]

RP SEQUENCE FROM N.A.
 CC TISSUE=Hepatopancreas;
 CC MEDLINE=95180724; PubMed=7875594;
 CC Zhang H., Okamoto N., Ikeda Y.;
 CC "Two c-myc genes from a tetraploid fish, the common carp (Cyprinus
 CC carpio).";
 CC Gene 153:231-236(1995).
 CC -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
 CC BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
 CC RECOGNIZES THE CORE SEQUENCE CAC[GA]TG. SEEMS TO ACTIVATE THE
 CC TRANSCRIPTION OF GROWTH-RELATED GENES (BY SIMILARITY)
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D37888; BAA07130.1; -
 CC HSSP; P01106; IA93.
 CC TRANSFAC; T03460; -
 CC InterPro; IPR001092; HLH_basic.
 CC InterPro; IPR003327; Myc-LZ.
 CC InterPro; IPR002418; TF_Myc.
 CC Pfam; PF00010; HLH; 1.
 CC Pfam; PF02344; Myc-LZ; 1.
 CC Pfam; PF01056; Myc-N-term; 1.
 CC PRINTS; PR00044; LEUZIPPRMYC.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HLH_1; FALSE_NEG.
 CC PROSITE; PS00888; HLH_2; 1.
 CC Nuclear protein; DNA-binding; Transcription regulation; Activator;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 198 206 POLY-SER.
 FT DOMAIN 207 228 POLY-GLU.
 FT DOMAIN 288 291 POLY-SER.
 FT DNA_BIND 317 330 BASIC DOMAIN.
 FT DOMAIN 331 370 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 376 397 LEUCINE-ZIPPER (POTENTIAL).
 FT CARBOHYD 58 58 O-LINKED (GLCNAC) (BY SIMILARITY).
 SQ SEQUENCE 401 AA; 45812 MW; 92501E4A4FCC75EF CRC64;
 Query Match 30.4%; Score 48; DB 1; Length 401;
 Best Local Similarity 58.3%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 16 LLSCHISSHQHD 27
 Db 256 LKRCHVSTHQHN 267
 | |||:||||
 RESULT 14
 MYC_BRARE
 ID MYC_BRARE STANDARD; PRT; 406 AA.
 AC P52160;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MYC protein (c-myc).
 GN MYC OR CMYC.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Danio.
 CC NCBI_TaxID=7955;


```

RN  SEQUENCE FROM N.A.
RX  MEDLINE=93233639; PubMed=8474440;
RA  Schreiber-Agus N., Horner J., Torres R., Chiu F.-C., DePinho R.A.;
RT  "Zebra fish myc family and max genes: differential expression and
RL  oncogenic activity throughout vertebrate evolution.";
RL  Mol. Cell. Biol. 13:2765-2775(1993).
CC  -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
CC  BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
CC  RECOGNIZES THE CORE SEQUENCE CAC[GA]TG. SEEMS TO ACTIVATE THE
CC  TRANSCRIPTION OF GROWTH-RELATED GENES (BY SIMILARITY).
CC  -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC  BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX (BY
CC  SIMILARITY).
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE KIDNEY, GILLS AND
CC  UTERUS.
CC  -!- DEVELOPMENTAL STAGE: FOUND IN LOW-ABUNDANCE IN THE TWO-CELL
CC  THROUGH EARLY SOMITE STAGES (<1.5 THROUGH 12 HRS) AND INCREASES
CC  DURING LATER STAGES OF GROWTH AND ORGAN DEVELOPMENT.
CC  -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC  TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L11710; AAA02482.1; -
DR  HSSP; P25912; 1HLO.
DR  TRANSFAC; T03450; -
DR  ZFIN; ZDB-GENE-990415-162; cmcy.
DR  InterPro; IPR001092; HLH_basic.
DR  InterPro; IPR003327; Myc-LZ.
DR  InterPro; IPR002418; TF_Myc.
DR  Pfam; PF000010; HLH; 1.
DR  Pfam; PF02344; Myc-LZ; 1.
DR  PRINTS; PR00044; LEUZIPPRMYC.
DR  SMART; SM00353; HLH; 1.
DR  PROSITE; PS00038; HLH_1; FALSE_NEG.
DR  PROSITE; PS00888; HLH_2; 1.
DR  Nuclear protein; DNA-binding; Transcription regulation; Activator;
KW  Glycoprotein.
FT  DOMAIN 199 207 POLY-SER.
FT  DOMAIN 208 229 POLY-GLU.
FT  DOMAIN 237 301 POLY-SER.
FT  DOMAIN 206 231 ASP/GLU-RICH (ACIDIC).
FT  DNA_BIND 322 335 BASIC DOMAIN.
FT  DOMAIN 326 337 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT  DOMAIN 336 375 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT  DOMAIN 381 402 LEUCINE-ZIPPER (POTENTIAL).
FT  CARBOHYD 58 58 O-LINKED (GLCNAC) (BY SIMILARITY).
SQ  SEQUENCE 406 AA; 46455 MW; 881D97A9B4B60E8A CRC64;

Query Match 30.4%; Score 48; DB 1; Length 406;
Best Local Similarity 58.3%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 LLSCHISSHQHD 27
DB 257 LKRVHSTHQHN 268

RESULT 15
MYC_ONCMY STANDARD; PRT; 414 AA.
AC P06646;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

```

DE  Myc protein (c-myc) (Fragment).
GN  MYC.
OS  Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX  NCBI_TaxID=8022;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86233301; PubMed=3520551;
RA  van Beneden R.J., Watson D.K., Chen T.T., Lautenberger J.A.;
RA  Papas T.S.;
RT  "Cellular myc (c-myc) in fish (rainbow trout): its relationship to
RT  other vertebrate myc genes and to the transforming genes of the MC29
RT  family of viruses.";
RL  Proc. Natl. Acad. Sci. U.S.A. 83:3698-3702(1986).
CC  -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
CC  BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
CC  RECOGNIZES THE CORE SEQUENCE CAC[GA]TG. SEEMS TO ACTIVATE THE
CC  TRANSCRIPTION OF GROWTH-RELATED GENES (BY SIMILARITY).
CC  -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC  BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX (BY
CC  SIMILARITY).
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC  TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M13048; AAA49604.1; -
DR  PIR; A25272; TVTRMC.
DR  HSSP; P25912; 1HLO.
DR  TRANSFAC; T03462; -
DR  InterPro; IPR001092; HLH_basic.
DR  InterPro; IPR003327; Myc-LZ.
DR  InterPro; IPR002418; TF_Myc.
DR  Pfam; PF000010; HLH; 1.
DR  Pfam; PF02344; Myc-LZ; 1.
DR  PRINTS; PR00044; LEUZIPPRMYC.
DR  SMART; SM00353; HLH; 1.
DR  PROSITE; PS00038; HLH_1; FALSE_NEG.
DR  PROSITE; PS00888; HLH_2; 1.
DR  Nuclear protein; DNA-binding; Transcription regulation; Activator.
KW  NON_TER 1 1
FT  DOMAIN 222 237 ASP/GLU-RICH (ACIDIC).
FT  DNA_BIND 324 337 BASIC DOMAIN.
FT  DOMAIN 326 337 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT  DOMAIN 338 377 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT  DOMAIN 383 404 LEUCINE-ZIPPER (POTENTIAL).
SQ  SEQUENCE 414 AA; 46642 MW; 06830232DB3D961A CRC64;

Query Match 30.4%; Score 48; DB 1; Length 414;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 LLSCHISSHQHD 27
DB 265 LKRVHSTHQHN 276

Search completed: October 21, 2003, 18:54:15
Job time : 10.25 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:42 ; Search time 44.75 seconds
(without alignments)
172.996 Million cell updates/sec

Title: FVIII_PHE328S

Perfect score: 158
Sequence: 1 TFLTAQTLMDLQGLLSCHISSHQHDGME 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTEMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_protist.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriaph.*

17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	144	91.1	355	11 Q8BQ43	Q8BQ43 mus musculus
2	143	90.5	2343	6 O62730	O62730 canis famil
3	143	90.5	2343	6 O18806	O18806 canis famil
4	75	47.5	782	4 O75659	O75659 homo sapien
5	75	47.5	1104	4 O75180	O75180 homo sapien
6	75	47.5	1158	4 Q9BQS7	Q9BQS7 homo sapien
7	75	47.5	1158	4 Q9C058	Q9C058 homo sapien
8	69	43.7	847	11 Q8C4S2	Q8C4S2 mus musculus
9	69	43.7	1157	11 Q9Z0Z4	Q9Z0Z4 mus musculus
10	66	40.5	594	3 Q96VT5	Q920H8 rattus norv
11	64	40.5	594	3 Q96VT5	Q96VT5 emericocella
12	62.5	39.6	1048	6 Q9XT27	Q9XT27 ovis aries
13	61	38.6	503	11 Q8BV37	Q8BV37 mus musculus
14	60	38.0	1084	11 Q9JL97	Q9JL97 rattus norv
15	59	37.3	669	2 Q8KKW5	Q8KKW5 rhizobium e
16	59	37.3	717	16 Q8XPV6	Q8XPV6 ralstonia s

Q9pa43 xylella fas
Q90zt2 brachydanio
Q989b7 rhizobium l
Q9uvv1 aspergillus
Q9kj88 marinomonas
Q49992 lycopersico
Q8bu82 mus musculus
Q8caf6 staphylococ
Q8ueq2 agrobacteri
Q9vxl1 drosophila
Q98bu9 rhizobium l
Q9cle4 pasteurella
Q8mtv0 schistosoma
Q67206 aquifex aeo
Q96wm9 botrytis ci
Q00292 aspergillus
Q93p73 microscilla
Q9kx4 bacillus ha
Q8xy1 ralstonia s
Q9ll75 arabidopsis
Q8mj65 loxodonta a
Q8k325 mus musculus
Q8k324 cricetus
Q8mj71 cuon alpinu
Q8mj64 chaetophrac
Q8mj68 oryctolagus
Q8mj69 plecotus au
Q8mj72 saguinus oe
Q8mj66 bos taurus

ALIGNMENTS

RESULT 1

Q8BQ43 ID Q8BQ43 PRELIMINARY; PRT; 355 AA.

AC Q8BQ43;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Coagulation factor VIII (Fragment).
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR ENBL; AK051579; BAC34681.1; -.
FT NON_TER 355
SQ SEQUENCE 355 AA; 40548 MW; 2D3C984EFA43F284 CRC64;

Query Match 91.1%; Score 144; DB 11; Length 355;
Best Local Similarity 90.0%; Pred. No. 2e-14;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLLSCHISSHQHDGME 30
|||||:|||||:|||||:|||||:|||||

Db 312 TFLTAQTLMDLQGLLSCHISSHQHDGME 341

RESULT 2

O62730 ID O62730 PRELIMINARY; PRT; 2343 AA.

AC O62730;
DT 01-AUG-1998 (TEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Factor VIII.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Spleen;
 RA Gordy P.W., Bowen R.A.;
 RT "Characterization of the canine factor VIII cDNA";
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 DE EMBL: AF049489; AAC05384.1; -;
 DE HSSP: P00451; ICFG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase_3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

 Query Match 90.5%; Score 143; DB 6; Length 2343;
 Best Local Similarity 90.0%; Pred. No. 2e-13;
 Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 QY 1 TELTAOTLLMDLQGLFLLSCHSHQHDGME 30
 DB 306 TELTAOTLLMDLQGLFLLSCHSHQHDGME 335

 RESULT 3
 O18806 PRELIMINARY; PRT; 2343 AA.
 AC O18806;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DE 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DE Factor VIII.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
 RT "The canine factor VIII cDNA and 5' flanking sequence";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 DE EMBL: AF016234; AB87412.1; -;
 DE HSSP: P00451; ICFG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase_3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 SQ SEQUENCE 2343 AA; 265829 MW; AB54FAE571C3B399 CRC64;

 Query Match 90.5%; Score 143; DB 6; Length 2343;
 Best Local Similarity 90.0%; Pred. No. 2e-13;
 Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 QY 1 TELTAOTLLMDLQGLFLLSCHSHQHDGME 30

Db 306 TELTAOTLLMDLQGLFLLSCHSHQHDGME 335

 RESULT 4
 O75659 PRELIMINARY; PRT; 782 AA.
 AC O75659;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DE 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE DJ46618.1 (Coagulation factor V (Activated protein C cofactor),
 DE coagulation factor VIII (Procoagulant component) and ceruloplasmin
 DE (BC 1.16.3.1, ferroxidase) like (fragment)).
 DE DJ46618.1.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkenson J.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DE EMBL: AL030998; CAA19742.1; -;
 DR HSSP: P00450; IKCW.
 DR InterPro: IPR001117; Cu-oxidase.
 DR Pfam: PF00394; Cu-oxidase_2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
 DE Coagulation.
 KW NON_TER 1
 FT NON_TER 782
 SQ SEQUENCE 782 AA; 87939 MW; 2FBD2824CC19D2AB CRC64;

 Query Match 47.5%; Score 75; DB 4; Length 782;
 Best Local Similarity 43.3%; Pred. No. 0.0037;
 Matches 13; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

 QY 1 TELTAOTLLMDLQGLFLLSCHSHQHDGME 30
 Db 257 TFTAEMVWPEPTWLIISQVNSHFQDGMQ 286

 RESULT 5
 O75180 PRELIMINARY; PRT; 1104 AA.
 AC O75180;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DE 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE KIAA0698 protein (Fragment).
 GN KIAA0698.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98403860; PubMed=9734811;
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro";
 RL DNA Res. 5:169-176(1998).
 DE EMBL: AB014598; BAA31673.2; -;
 DR HSSP: P00450; IKCW.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; MultiCu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase_3.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DE PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 FT NON_TER 1

```

DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; Multicu_oxidase2.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1158 AA; 130448 MW; CD032199E2E2868D CRC64;

Query Match 47.5%; Score 75; DB 4; Length 1158;
Best Local Similarity 43.3%; Pred. No. 0.0055;
Matches 13; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGGFLLSCHISSHQHDGME 30
   ||::: : : : : ||::: ||:::
Db 329 TFTVAEMVFPWEPGTWLISCQVNSHFRDGMQ 358

RESULT 8
Q8C4S2 PRELIMINARY; PRT; 847 AA.
ID Q8C4S2
AC Q8C4S2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HEPHAESTIN homolog.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
FR EMBL; AK081330; BAC38197.1; -.
SQ SEQUENCE 847 AA; 95136 MW; 778CBE0617243402 CRC64;

Query Match 43.7%; Score 69; DB 11; Length 847;
Best Local Similarity 40.0%; Pred. No. 0.036;
Matches 12; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGGFLLSCHISSHQHDGME 30
   ||::: : : : : ||::: ||:::
Db 329 TFTVAEMVFPKSGTWLISCVEVNSHLRSGMQ 358

RESULT 9
Q9Z0Z4 PRELIMINARY; PRT; 1157 AA.
ID Q9Z0Z4
AC Q9Z0Z4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hephaestin.
DE Mus musculus (Mouse).
GN HEPH.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99140771; PubMed=9988272;
RA Vulpe C.D.; Kuo Y.M.; Murphy T.L., Cowley L., Askwith C., Libina N.,
RA Gitschier J., Anderson G.J.;
RT "Hephaestin, a ceruloplasmin homologue implicated in intestinal iron
RT transport, is defective in the sla mouse.";
RL Nat. Genet. 21:195-199(1999).
FR EMBL; AF082567; RAD16035.1; -.
DB HSSP; P00450; 1KCW.

```

```

DR MGD; MGI:1332240; Heph.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidase2.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1157 AA; 129682 MW; 411BB1DA2BFEC2FD CRC64;

Query Match 43.7%; Score 69; DB 11; Length 1157;
Best Local Similarity 40.0%; Pred. No. 0.049;
Matches 12; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||: : : : : |||: |||
Db 329 TFTAEMVVPQKSGTWLISCEVNSHLRSGMQ 358

RESULT 10
Q920H8 PRELIMINARY; PRT; 1157 AA.
ID Q920H8;
AC Q920H8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hephæstín.
GN HEPH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Frazer D.M., Vulpe C.D., McKie A.T., Wilkins S.J., Trinder D.,
RA Clegghorn G.J., Anderson G.J.;
RT "Cloning and gastrointestinal expression of rat hephaestin:
RT relationship to other iron transport proteins.";
RL Am. J. Physiol. 281:G931-G939(2001).
DR EMBL; AF246120; AAL08217.1; -.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidase2.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1157 AA; 129593 MW; 0C626FA3E2F51DE2 CRC64;

Query Match 41.8%; Score 66; DB 11; Length 1157;
Best Local Similarity 36.7%; Pred. No. 0.15;
Matches 11; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGGQFLLSCHSHQHDGME 30
|||: : : : : |||: |||
Db 1027 TFEVVENVASNPGAWLHCHVTDHVGAGME 1056

RESULT 11
Q96VT5 PRELIMINARY; PRT; 594 AA.
ID Q96VT5;
AC Q96VT5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Laccase precursor (EC 1.10.3.2).
GN TIIA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RA Scherer M.;
RL Thesis (2001), Department of Microbiology,
RL Philipps-Universitaet-Marburg, Marburg, Germany.

```

```

DR EMBL; AJ305224; CAC59820.1; -.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidase2.
DR Pfam: PF00394; Cu-oxidase; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Oxidoreductase; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 594 LACCASE.
SQ SEQUENCE 594 AA; 65816 MW; 16A1E82F24F7F01C CRC64;

Query Match 40.5%; Score 64; DB 3; Length 594;
Best Local Similarity 44.8%; Pred. No. 0.15;
Matches 13; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGGQFLLSCHSHQHDGM 29
|||: : : : : |||: |||
Db 545 TWIALRYVNVNPGFLLHCHIQHSGGM 573

RESULT 12
Q9XT27 PRELIMINARY; PRT; 1048 AA.
ID Q9XT27;
AC Q9XT27;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ceruloplasmin.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99384006; PubMed=10452945;
RA Lockhart P.J., Mercer J.F.B.;
RT "Cloning and expression analysis of the sheep ceruloplasmin cDNA.";
RL Gene 236:251-257(1999).
DR EMBL; AF134814; AAD41477.1; -.
DR HSSP; P00450; 1KCW.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidase2.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1048 AA; 119125 MW; 925F16D7B0549CBB CRC64;

Query Match 39.8%; Score 62.5; DB 6; Length 1048;
Best Local Similarity 50.0%; Pred. No. 0.48;
Matches 15; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

QY 4 TAQTLLM--DLGGQFLLSCHSHQHDGME 30
|||: : : : : |||: |||
Db 1005 TYQLEMTPTPGIWLHCHVTDHVGAGME 1034

RESULT 13
Q8BV37 PRELIMINARY; PRT; 503 AA.
ID Q8BV37;
AC Q8BV37;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to GPI-anchored ceruloplasmin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;

```

```

RX MEDLINE-22354683; PubMed-12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK080701; BAC37984.1; -.
FT NON_TER 1 1
SQ SEQUENCE 503 AA; 58067 MW; 042331F60B1853EB CRC64;

Query Match 38.0%; Score 61; DB 11; Length 503;
Best Local Similarity 37.9%; Pred. No. 0.39;
Matches 11; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLLSCHISSHQHDGM 29
: : : : : : : : : : : : : : : :
Db 459 TYQLEMFPTPGTWLLHCHVTDHVGAM 487

RESULT 14
Q9JL97 PRELIMINARY; PRT; 1084 AA.
AC Q9JL97;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GPI-anchored ceruloplasmin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague Dawley; TISSUE=Brain;
RX MEDLINE-20127919; PubMed=10660599;
RA Patel B.N., Dunn R.J., David S.;
RT "Alternative RNA Splicing Generates a Glycosylphosphatidylinositol-
RT anchored Form of Ceruloplasmin in Mammalian Brain.";
RL J. Biol. Chem. 275:4305-4310(2000).
DR EMBL: AF202115; AAF34175.1; -.
DR HSSP; P00450; IKCW.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidse2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1084 AA; 123749 MW; 95D65B4154257C55 CRC64;

Query Match 38.0%; Score 60; DB 11; Length 1084;
Best Local Similarity 37.9%; Pred. No. 1.2;
Matches 11; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLLSCHISSHQHDGM 29
: : : : : : : : : : : : : : : :
Db 1016 TYQLEMFPTPGTWLLHCHVTDHVGAM 1044

RESULT 15
Q8KKW5 PRELIMINARY; PRT; 669 AA.
AC Q8KKW5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN yP076.
OS Rhizobium etli.
OC Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-CFN42;
RX MEDLINE-91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
RT leguminosarum bv. phaseoli.";
RL J. Bacteriol. 173:2411-2419(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CFN42;
RX MEDLINE-974119521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
RT of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CFN42;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase and RecA are required to exert
RT incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80928; AAM55015.1; -.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu_oxidse2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 669 AA; 73046 MW; CE98B723C7B72996 CRC64;

Query Match 37.3%; Score 59; DB 2; Length 669;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 12 LGQFLSCHISSHQHDGM 29
: : : : : : : : : : : :
Db 632 IGEFVLHCHTLDHEDQGM 649

Search completed: October 21, 2003, 18:57:22
Job time : 45.75 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:44:27 ; Search time 49.5 Seconds
(without alignments)
96.198 Million cell updates/sec

Title: FVIII_PHE328X

Perfect score: 155

Sequence: 1 TFLTAQLLLMDLQGLLXCHISSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155	100.0	372	16 AAR73019	Human Factor-VIII
2	155	100.0	400	16 AAR67709	B-domain deleted F
3	155	100.0	720	16 AAR74088	Factor-VIII heavy
4	155	100.0	729	16 AAR74089	Factor-VIII heavy
5	155	100.0	740	16 AAR76959	Human Factor-VIII/
6	155	100.0	740	16 AAR76961	Human Factor-VIII
7	155	100.0	740	16 AAR76962	Human Factor-VIII
8	155	100.0	740	16 AAR73021	Human Factor-VIII
9	155	100.0	740	16 AAR74090	Factor-VIII heavy

10	155	100.0	740	16 AAR76982	Human factor VIII
11	155	100.0	1014	8 AAP71139	Factor VIII:c varia
12	155	100.0	1383	18 AAW33227	Procoagulant-activ
13	155	100.0	1383	18 AAW33228	Procoagulant-activ
14	155	100.0	1383	18 AAW33229	Procoagulant-activ
15	155	100.0	1424	9 AAP80268	Modified factor VI
16	155	100.0	1424	10 AAP91169	Sequence of 740 Ar
17	155	100.0	1424	22 AAB48842	Mutant mature huma
18	155	100.0	1424	23 AAP18622	Human mature B-dom
19	155	100.0	1425	9 AAP80267	Modified factor VI
20	155	100.0	1438	21 AAB01262	B-domain deleted f
21	155	100.0	1440	12 AAR12971	Factor VIII:SQ. U
22	155	100.0	1445	23 ABG92540	IE B-domain-delete
23	155	100.0	1447	23 ABG92541	5Arg B-domain-dele
24	155	100.0	1457	19 AAW46246	Human factor VIII
25	155	100.0	1457	19 AAW44372	Human Factor VIII
26	155	100.0	1457	20 AAY21675	Beta-domain delete
27	155	100.0	1459	22 AAE10827	Human factor VIII
28	155	100.0	1459	22 AAE10832	Human factor VIII
29	155	100.0	1459	22 AAE10833	Human factor VIII
30	155	100.0	1471	18 AAW23414	Human B-domain del
31	155	100.0	1471	22 AAW67959	Amino acid sequenc
32	155	100.0	1516	9 AAP80265	Modified factor VI
33	155	100.0	1661	18 AAW18670	Factor VIII-dB695-
34	155	100.0	2098	17 AAR86863	Factor-VIII. Homo
35	155	100.0	2332	8 AAP71726	Factor VIII:c varia
36	155	100.0	2332	8 AAP71727	Factor VIII:c varia
37	155	100.0	2332	8 AAP71728	Factor VIII:c varia
38	155	100.0	2332	8 AAP71729	Factor VIII:c varia
39	155	100.0	2332	14 AAR43257	Human Factor VIII.
40	155	100.0	2332	18 AAW33222	Procoagulant-activ
41	155	100.0	2332	18 AAW33223	Procoagulant-activ
42	155	100.0	2332	18 AAW33224	Procoagulant-activ
43	155	100.0	2332	18 AAW33225	Procoagulant-activ
44	155	100.0	2332	18 AAW33226	Procoagulant-activ
45	155	100.0	2332	19 AAW53483	Human factor VIII.

ALIGNMENTS

RESULT 1
AAR73019
ID AAR73019 standard; peptide: 372 AA.
AC AAR73019;
XX
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
DE Human Factor-VIII N-terminal fragment.
XX
XX Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX Homo sapiens.
XX
XX WO9513301-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK00424.
XX
PR 12-NOV-1993; 93DK-0001281.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
PI Persson E;
XX
XX WPI; 1995-194038/25.
XX
PT Crosslinked Factor VIII polypeptide which is stable - is prepd. using
PT bis(sulphosuccinimidy) suberate or disuccinimydyl suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting

PT activity
 PS Disclosure; Page 18; 36pp; English.
 XX
 CC This is the N-terminal fragment of human Factor-VIII which
 CC may be crosslinked resulting in increased stability and retention
 CC of high activity over extended periods of time after activation by
 CC thrombin. The polypeptide is used to prevent or treat diseases
 CC caused by the absence or deficiency of Factor-VIII in a subject
 CC such as haemophilia.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 372 AA;
 SQ
 Query Match 100.0%; Score 155; DB 16; Length 372;
 Best Local Similarity 96.7%; Pred. No. 2e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLGQFLXCHISSHQHDGME 30
 DB 292 TELTAQTLLMDLGQFLXCHISSHQHDGME 321
 RESULT 2
 AAR67709
 ID AAR67709 standard; Protein; 400 AA.
 XX
 AC AAR67709;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-JUL-1995 (first entry)
 XX
 DE B-domain deleted Factor-VIII.
 XX
 KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;
 KW adenovirus; vector.
 XX
 OS Homo sapiens.
 XX
 PN WO9429471-A1.
 XX
 XX 22-DEC-1994.
 XX
 PF 13-APR-1994; 94WO-US04075.
 XX
 PR 10-JUN-1993; 93US-0074920.
 PR 25-MAR-1994; 94US-0218335.
 XX
 PA (GENE-) GENETIC THERAPY INC.
 XX
 PI Connelly S, Kaleko M, Smith T;
 XX
 DR WPI; 1995-036495/05.
 DR N-PSDB; AAQ76016.
 XX
 PT New adenoviral vectors for treatment of haemophilia - contg. a
 PT DNA sequence encoding a clotting factor, partic. Factor VIII or
 PT Factor IX
 XX
 PS Disclosure; Fig. 17A-17C; 116pp; English.
 XX
 CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
 CC used to construct recombinant adenovirus vectors that produced
 CC therapeutic levels of the clotting factor when administered to an
 CC animal host, potentially providing hemophilia A gene therapy.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 400 AA;
 SQ
 Query Match 100.0%; Score 155; DB 16; Length 400;
 Best Local Similarity 96.7%; Pred. No. 2.1e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLXCHISSHQHDGME 30
 DB 310 TELTAQTLLMDLGQFLXCHISSHQHDGME 339
 RESULT 3
 AAR74088
 ID AAR74088 standard; protein; 720 AA.
 XX
 AC AAR74088;
 XX
 DT 25-MAR-2003 (updated)
 DT 04-NOV-1995 (first entry)
 XX
 DE Factor-VIII heavy chain N-terminal fragment.
 XX
 KW human; Factor VIII; heavy chain; N-terminal fragment;
 KW thrombin cleavage; blood-clotting.
 XX
 OS Homo sapiens.
 XX
 PN WO9513300-A1.
 XX
 PD 18-MAY-1995.
 XX
 PF 10-NOV-1994; 94WO-DK00423.
 XX
 PR 12-NOV-1993; 93DK-0001280.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Ezban Rasmussen M, Kjalke M;
 XX
 DR WPI; 1995-194037/25.
 XX
 PT Factor VIII polypeptide(s) comprising a heavy chain shorter than native
 PT Al-A2 domain - are easier to produce recombinantly and retain coagulant
 PT activity, may be used to treat patients who have developed antibodies to
 PT C-terminal epitope(s) of Factor VIII
 XX
 PS Claim 3; Page 24-26; 51pp; English.
 XX
 CC The sequence represents N-terminal residues 1-720 of a human Factor-
 CC VIII heavy chain. The sequence is shorter than the Al-A2 domain,
 CC and is produced by treating a polypeptide containing the full Al-A2
 CC domain of full-length Factor-VIII with a protease, e.g. thrombin.
 CC The fragment has the same specific activity as full-length
 CC Factor-VIII in a chromogenic assay, but activity is a factor of two
 CC lower in a clotting assay, and the fragment is activated by thrombin
 CC at a slower rate and to a lower level than fragments 1-740 (AAR74090),
 CC 1-729 (AAR74089) and plasma Factor-VIII. The fragment may be produced
 CC recombinantly to reduce production costs and improve safety, and
 CC production levels and stability are higher than for the full-length
 CC form. The fragment may be used to treat patients who have developed
 CC antibodies against epitopes in the C-terminal part of the heavy chain.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 720 AA;
 SQ
 Query Match 100.0%; Score 155; DB 16; Length 720;
 Best Local Similarity 96.7%; Pred. No. 4.1e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLGQFLXCHISSHQHDGME 30
 DB 292 TELTAQTLLMDLGQFLXCHISSHQHDGME 321
 RESULT 4
 AAR74089
 ID AAR74089 standard; protein; 729 AA.
 XX
 AC AAR74089;

```

XX 25-MAR-2003 (updated)
DT 04-NOV-1995 (first entry)
DE Factor-VIII heavy chain N-terminal fragment.
XX human; Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX Homo sapiens.
XX WO9513300-A1.
XX 18-MAY-1995.
XX 10-NOV-1994; 94WO-DK00423.
XX 12-NOV-1993; 93DK-0001280.
XX (NOVO ) NOVO-NORDISK AS.
XX Ezban Rasmussen M, Kjalke M;
XX WPI; 1995-194037/25.
XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
PT Al-A2 domain - are easier to produce recombinantly and retain coagulant
PT activity, may be used to treat patients who have developed antibodies to
PT C-terminal epitope(s) of Factor VIII
XX Claim 2; Page 27-29; 51pp; English.
XX The sequence represents N-terminal residues 1-729 of a human Factor-
CC VIII heavy chain. The sequence is shorter than the Al-A2 domain,
CC and is produced by treating a polypeptide containing the full Al-A2
CC domain of full-length Factor-VIII with a protease, e.g. thrombin.
CC The fragment has the same coagulant specific activity as full-length
CC Factor-VIII in a chromogenic assay, and is activated by thrombin at
CC a similar rate. The fragment may be produced recombinantly to reduce
CC production costs and improve safety, and production levels and
CC stability are higher than for the full-length form. The fragment may
CC be used to treat patients who have developed antibodies against
CC epitopes in the C-terminal part of the heavy chain.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 729 AA;
SQ Query Match 100.0%; Score 155; DB 16; Length 729;
Best Local Similarity 96.7%; Pred. No. 4.1e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
Db |||||
292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

RESULT 5
AAR76959
ID AAR76959 standard; protein; 740 AA.
AC AAR76959;
XX
XX 25-MAR-2003 (updated)
DT 09-MAR-1996 (first entry)
XX Human Factor-VIII/Factor-VIIIa derivative.
DE Human Factor-VIII derivative.
XX Factor-VIII; therapeutic; blood-clotting.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 403

```

```

FT /label= Asp absent or Asn, Ser, Thr, Ala
FT Misc-difference 404
FT /label= Asp absent or Asn, Ser, Thr, Ala
FT Misc-difference 433
FT /label= Asp absent or Asn, Ser, Thr, Ala
FT Misc-difference 482
FT /label= Asp absent or Asn, Ser, Thr, Ala
FT Misc-difference 500
FT /label= Asp absent or Asn, Ser, Thr, Ala
FT Misc-difference 434
FT /label= Glu absent or Gln, Ser, Thr, Ala
FT Misc-difference 440
FT /label= Glu absent or Gln, Ser, Thr, Ala
XX WO9518829-A1.
XX 13-JUL-1995.
XX 06-JAN-1995; 95WO-DK00010.
XX 07-JAN-1994; 94DK-0000030.
XX (NOVO ) NOVO-NORDISK AS.
XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX WPI; 1995-255041/33.
XX Novel factor VIII derivative used to treat haemophilia - and
PT comprises a functional A2 domain containing a mutation at one or
PT more Cys residues.
XX Disclosure; Page 11-14; 24pp; English.
XX The new Factor-VIII/Factor-VIIIa derivative comprises a functional
CC A2 domain in which amino acid residues, as indicated in the
CC Features, are deleted or substituted by another amino acid so as to
CC increase the overall positive charge. Asp-403, Asp-404, Asp-433,
CC Asp-482 and Asp-500 are preferably substituted by Asn; Glu-434 and
CC Glu-440 are preferably substituted by Gln. The new derivative has
CC the same activity as the wild-type Factor-VIII but with improved
CC stability (the activity is maintained for a longer period compared
CC to the rapid decline of the activity of wt factor-VIII). The new
CC derivative can be used in a composition for treating diseases caused
CC by an absence or deficiency of Factor-VIII, especially haemophilia.
CC (Updated on 25-MAR-2003 to correct DR field.)
XX Sequence 740 AA;
SQ Query Match 100.0%; Score 155; DB 16; Length 740;
Best Local Similarity 96.7%; Pred. No. 4.2e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
Db |||||
292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

RESULT 6
AAR76961
ID AAR76961 standard; protein; 740 AA.
XX
XX AAR76961;
XX 09-MAR-1996 (first entry)
DT Human Factor-VIII derivative.
XX Factor-VIII; therapeutic; blood-clotting.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 403

```


FT Misc-difference 692 /label= absent or Ala, Thr, Ser, Gly or Asp
 FT Misc-difference 720 /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729 /label= absent or Val, Ala or Ile
 XX WO9518827-Al.
 PN
 XX 13-JUL-1995.
 PD
 XX
 XX 06-JAN-1995; 95WO-DK000008.
 PF
 XX 07-JAN-1994; 94DK-0000032.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
 PI WPI; 1995-255039/33.
 DR
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX Disclosure; Page 14-17; 30pp; English.
 PS
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is deleted or replaced with another amino acid
 CC residue, preferably Ser (see AAR76962). Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX
 SQ Sequence 740 AA;
 Query Match 100.0%; Score 155; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 4.2e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLGQFLLCXCHSHQHDGME 30
 DB 292 TELTAQTLLMDLGQFLLCXCHSHQHDGME 321
 RESULT 7
 AAR76962
 ID AAR76962 standard; protein; 740 AA.
 XX
 AC AAR76962;
 XX
 DT 09-MAR-1996 (first entry)
 XX
 DE Human Factor-VIII derivative.
 XX
 KW Factor-VIII; therapeutic; blood-clotting.
 XX
 OS Homo sapiens.
 XX
 EH Key Location/Qualifiers
 FT Misc-difference 692 /label= Cys substituted by Ser
 FT Misc-difference 720 /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729 /label= absent or Val, Ala or Ile
 XX WO9518827-Al.
 PN
 XX

PD 13-JUL-1995.
 XX
 XX 06-JAN-1995; 95WO-DK000008.
 PF
 XX 07-JAN-1994; 94DK-0000032.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
 PI WPI; 1995-255039/33.
 DR
 XX Novel factor VIII derivative used to treat haemophilia - and
 PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX Disclosure; Page 18-20; 30pp; English.
 PS
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is replaced with Ser. For other (less preferred)
 CC substitutions at this site, see AAR76961. Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX
 SQ Sequence 740 AA;
 Query Match 100.0%; Score 155; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 4.2e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TELTAQTLLMDLGQFLLCXCHSHQHDGME 30
 DB 292 TELTAQTLLMDLGQFLLCXCHSHQHDGME 321
 RESULT 8
 AAR73021
 ID AAR73021 standard; peptide; 740 AA.
 XX
 AC AAR73021;
 XX
 DT 25-MAR-2003 (updated)
 XX
 DT 21-NOV-1995 (first entry)
 XX
 DE Human Factor-VIII N-terminal fragment.
 XX
 KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
 XX
 OS Homo sapiens.
 XX
 PN WO9513301-Al.
 XX
 PD 18-MAY-1995.
 XX
 PF 10-NOV-1994; 94WO-DK00424.
 XX
 PR 12-NOV-1993; 93DK-0001281.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Persson E;
 PI
 XX WPI; 1995-194038/25.
 DR
 XX Crosslinked Factor VIII polypeptide which is stable - is prepd. using
 PT bis(sulphosuccinimidy) suberate or disuccinimetyl suberate in the
 PT presence of polysorbate 80 to produce a coagulant with long lasting
 PT activity

```

XX PS Disclosure; Page 21; 36pp; English.
XX CC This is the N-terminal fragment of human Factor-VIII which may be
CC crosslinked resulting in increased stability and retention of high
CC activity over extended periods of time after activation by thrombin.
CC The polypeptide is used to prevent or treat diseases caused by the
CC absence or deficiency of Factor-VIII in a subject such as
CC haemophilia.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 740 AA;

Query Match 100.0%; Score 155; DB 16; Length 740;
Best Local Similarity 96.7%; Pred. No. 4.2e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TFLTAQTLLMDLGGFLXCHISSHQHDCME 30
   |||||
Db 292 TFLTAQTLLMDLGGFLXCHISSHQHDCME 321

RESULT 9
AAR74090
ID AAR74090 standard; protein; 740 AA.
XX AC AAR74090;
XX DT 25-MAR-2003 (updated)
XX DT 04-NOV-1995 (first entry)
XX DE Factor-VIII heavy chain N-terminal fragment.
XX KW human; Factor VIII; heavy chain; N-terminal fragment;
XX KW thrombin cleavage; blood-clotting.
XX OS Homo sapiens.
XX PN WO9513300-A1.
XX PD 18-MAY-1995.
XX PF 10-NOV-1994; 94WO-DK00423.
XX PR 12-NOV-1993; 93DK-0001280.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Ezban Rasmussen M, Kjalke M;
XX PS WPI; 1995-194037/25.
XX DR
XX PT Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
XX PT activity, may be used to treat patients who have developed antibodies to
XX PT C-terminal epitope(s) of Factor VIII
XX PS Disclosure; Page 30-32; 51pp; English.
XX CC The sequence represents N-terminal residues 1-740 of a human Factor-
XX CC VIII heavy chain. The sequence contains entire A1 and A2 domains,
XX CC and truncated forms (1-720 in (AAR74088) and 1-729 in (AAR74090))
XX CC may be produced by treatment with a protease, e.g. thrombin. The
XX CC C-terminally truncated fragments have the same coagulant specific
XX CC activity as full-length Factor-VIII, and may be produced
XX CC recombinantly to reduce production costs and improve safety, giving
XX CC higher production levels and stability than for the full-length form.
XX CC The fragments may be used to treat patients who have developed
XX CC antibodies against epitopes in the C-terminal part of the heavy chain.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 740 AA;

Query Match 100.0%; Score 155; DB 16; Length 740;
Best Local Similarity 96.7%; Pred. No. 4.2e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TFLTAQTLLMDLGGFLXCHISSHQHDCME 30
   |||||
Db 292 TFLTAQTLLMDLGGFLXCHISSHQHDCME 321

RESULT 9
AAR74090
ID AAR74090 standard; protein; 740 AA.
XX AC AAR74090;
XX DT 25-MAR-2003 (updated)
XX DT 04-NOV-1995 (first entry)
XX DE Factor-VIII heavy chain N-terminal fragment.
XX KW human; Factor VIII; heavy chain; N-terminal fragment;
XX KW thrombin cleavage; blood-clotting.
XX OS Homo sapiens.
XX PN WO9513300-A1.
XX PD 18-MAY-1995.
XX PF 10-NOV-1994; 94WO-DK00423.
XX PR 12-NOV-1993; 93DK-0001280.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Ezban Rasmussen M, Kjalke M;
XX PS WPI; 1995-194037/25.
XX DR
XX PT Factor VIII polypeptide(s) comprising a heavy chain shorter than native
XX PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
XX PT activity, may be used to treat patients who have developed antibodies to
XX PT C-terminal epitope(s) of Factor VIII
XX PS Disclosure; Page 30-32; 51pp; English.
XX CC The sequence represents N-terminal residues 1-740 of a human Factor-
XX CC VIII heavy chain. The sequence contains entire A1 and A2 domains,
XX CC and truncated forms (1-720 in (AAR74088) and 1-729 in (AAR74090))
XX CC may be produced by treatment with a protease, e.g. thrombin. The
XX CC C-terminally truncated fragments have the same coagulant specific
XX CC activity as full-length Factor-VIII, and may be produced
XX CC recombinantly to reduce production costs and improve safety, giving
XX CC higher production levels and stability than for the full-length form.
XX CC The fragments may be used to treat patients who have developed
XX CC antibodies against epitopes in the C-terminal part of the heavy chain.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 740 AA;

Query Match 100.0%; Score 155; DB 16; Length 740;
Best Local Similarity 96.7%; Pred. No. 4.2e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TFLTAQTLLMDLGGFLXCHISSHQHDCME 30
   |||||
Db 292 TFLTAQTLLMDLGGFLXCHISSHQHDCME 321

RESULT 10
AAR76982
ID AAR76982 standard; protein; 740 AA.
XX AC AAR76982;
XX DT 21-FEB-1996 (first entry)
XX DE Human factor VIII A2-domain derivative.
XX KW Factor VIII; human; haemophilia; thrombin; protein C; plasmin;
XX KW serine protease; recombination; therapy; deficiency.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 720 /label= Gln, Ser, Thr, Val, Ala
XX FT Misc-difference 729 /label= Val, Ala, Ile
XX PN WO9518828-A1.
XX PD 13-JUL-1995.
XX PF 06-JAN-1995; 95WO-DK00009.
XX PR 07-JAN-1994; 94DK-0000031.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX PS WPI; 1995-255040/33.
XX DR
XX PT Novel factor VIII derivative with resistance to enzymatic cleavage -
XX PT and comprises a functional A2 domain where Glu720 and/or Tyr729 is
XX PT deleted/substituted.
XX PS Claim 1; Page 11-14; 26pp; English.
XX CC This sequence represents the A2 domain of a human factor VIII derivative.
XX CC Factor VIII is a large glycoprotein which is present in plasma at low
XX CC concentrations. Factor VIII is an essential part of the clotting
XX CC reaction in response to a wound. Factor VIII is susceptible to cleavage
XX CC by thrombin, activated protein C, plasmin, and other serine proteases.
XX CC Full length factor VIII consists of three repeats of the A-domain, a
XX CC B-domain and 2 repeats of the C-domain. Active factor VIII has the A1
XX CC domain cleaved off. Factor VIII is too unstable for use in recombinant
XX CC techniques. Factor VIII containing this sequence has improved stability
XX CC and shows resistance against enzymatic activity present in mammalian
XX CC cells. This means that factor VIII containing this sequence can be used
XX CC in recombinant techniques. A factor VIII derivative can be used for
XX CC treating diseases caused by an absence or deficiency of factor VIII (in
XX CC the same way as normal factor VIII) e.g. haemophilia. The advantage with
XX CC using a recombinant factor VIII also includes no need for lots of donors
XX CC in order to get a sufficient amount. Also, there is no long purification
XX CC process, and there is no risk of transmission of blood-borne diseases
XX CC such as HIV.
XX SQ Sequence 740 AA;

Query Match 100.0%; Score 155; DB 16; Length 740;
Best Local Similarity 96.7%; Pred. No. 4.2e-14;

```

```

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 11
AAP71139
ID AAP71139 standard; Protein; 1014 AA.
XX
AC AAP71139;
XX
DT 25-MAR-2003 (updated)
DT 14-MAR-1991 (first entry)
XX
DE Factor VIII:c variant.
XX
KW Factor VIII:c; variant; proteolysis; resistance;
KW pro coagulation activity.
XX
OS Homo sapiens.
XX
PN WO8707144-A.
XX
PD 03-DEC-1987.
XX
PF 29-MAY-1987; 87WO-US01299.
XX
PR 29-MAY-1986; 86US-0868410.
PR 18-NOV-1986; 86US-0932767.
PR 09-DEC-1986; 86US-0939658.
XX
PA (GEM) ) GENETICS INST INC.
XX
PI Kaufman RJ, Pittman DD, Toole JJ;
XX
WPI; 1987-348539/49.
XX
PT New deletion and replacement variants of factor VIII:c - resistant
PT to proteolysis but retaining pro coagulant activity, and mew DNA
PT coding sequences.
XX
PS Disclosure; Page 1; 42pp; English.
XX
CC The full-length human factor VIII:c cDNA has been set forth in
CC detail in WO8501961. This sequences is an example of the formula
CC A-X-B, wherein A-Ala1-Arg372, B-Ser1690-Tyr2332 and X=0-1316
CC amino acids substantially duplicative of sequences of amino acids
CC within the sequence Arg372-Ser1690 of the full-length sequence.
CC Here X=0; producing a Arg371-Ser1690 fusion protein.
CC One or more deletions or replacements at Arg 220, 226, 279, 282,
CC 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
CC variants which are more resistant to specific proteolytic cleavage
CC compared with natural factor VIII:c. Pro-coagulant activity and
CC thrombin activatability have been retained.
CC See also AAP71726-29.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1014 AA;

Query Match 100.0%; Score 155; DB 8; Length 1014;
Best Local Similarity 96.7%; Pred. No. 5.9e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 12
AAW33227
ID AAW33227 standard; protein; 1383 AA.

```

```

XX AAW33227;
XX 30-APR-1998 (first entry)
XX Procoagulant-active human factor VIII:c (FVIII) mutant protein.
XX
XX Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
XX Synthetic.
OS Homo sapiens.
XX
FH Key
FT Region 1..346
FT /note= "factor VIIIA heavy chain"
FT Region 741..1383
FT /note= "factor VIIIA light chain"
FT Domain 1..329
FT /note= "A1 domain"
FT Domain 1..179
FT /note= "plastocyanin-like domain 1"
FT Domain 187..329
FT /note= "plastocyanin-like domain 2"
FT Domain 380..711
FT /note= "A2 domain"
FT Misc_feature 711..746
FT /note= "a spacer of the sequence
FT SFQNSRHPSTROKQFNATIPENDIEKTPWF
FT AHRTMPKIQNVSSDILMLL is inserted
FT between domains A2 and A3"
FT
FT Domain 380..554
FT /note= "plastocyanin-like domain 3"
FT Domain 564..711
FT /note= "plastocyanin-like domain 4"
FT Domain 746..1073
FT /note= "A3 domain"
FT Domain 1073..1221
FT /note= "C1 domain"
FT Domain 1226..1378
FT /note= "C2 domain"
FT Cleavage-site 372..373
FT /note= "by thrombin"
FT Disulfide-bond 153..179
FT /note= "probable"
FT Disulfide-bond 528..554
FT /note= "probable"
FT Misc-difference 740
FT /label= R740A
FT /note= "wild type Arg replaced with Ala"
XX
XX WO9740145-A1.
XX
XX 30-OCT-1997.
XX
XX 24-APR-1997; 97WO-US06563.
XX
XX 15-MAY-1996; 96US-0017785.
XX 24-APR-1996; 96US-0016117.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Amano K, Kaufman RJ, Pipe SW;
XX
XX WPI; 1997-535830/49.
XX
XX Modified human pro-coagulant active factor VIII - can be
XX administered to haemophiliacs, i.e. factor VIII replacement therapy
XX
XX Claim 20; Page -: 57pp; English.
PS

```

XX The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 CC von Willebrand factor binding site, a mutation at Arg740 and an addition
 CC of an amino acid sequence spacer between the A2 and A3 domains. Factor
 CC VIII, along with calcium and phospholipid, acts as a cofactor for factor
 CC IXA, when it converts factor X to the activated form (factor XA). FVIII
 CC is the coagulation factor deficient in the X-chromosome-linked bleeding
 CC disorder haemophilia A. Several other mutant FVIII proteins have also
 CC been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
 CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII
 CC mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
 CC protein C (APC) cleavage. The present FVIII mutant can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

XX Sequence 1383 AA;

Query Match 100.0%; Score 155; DB 18; Length 1383;
 Best Local Similarity 96.7%; Pred. No. 8.4e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
 |||||
 Db 292 TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 13

AAW33228
 ID AAW33228 standard; protein; 1383 AA.

XX AC AAW33228;

XX DT 30-APR-1998 (first entry)

XX DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.

XX KW Pro-coagulant active factor VIII: FVIII; haemophilia A;

XX KW recombinant secretion; pro-coagulant activity; resistance;

XX KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;

XX KW von Willebrand factor binding site; binding affinity;

XX KW FVIII replacement therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Region 1..346

FT FT /note= "factor VIIIA heavy chain"

FT Region 741..1383

FT FT /note= "factor VIIIA light chain"

FT Domain 1..329

FT FT /note= "A1 domain"

FT Domain 1..179

FT FT /note= "plastocyanin-like domain 1"

FT Domain 187..329

FT FT /note= "plastocyanin-like domain 2"

FT Domain 380..711

FT FT /note= "A2 domain"

FT Misc_feature 711..746

FT FT /note= "a spacer of the sequence

FT FT SFSQNSRHPSTRQKFNATIPENDIEKTPWF

FT FT AHRTPMKIQNVSSDLMLL is inserted

FT FT between domains A2 and A3"

FT Domain 380..554

FT FT /note= "plastocyanin-like domain 3"

FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 746..1073
 FT /note= "A3 domain"
 FT Domain 1073..1221
 FT /note= "C1 domain"
 FT Domain 1226..1378
 FT /note= "C2 domain"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Misc-difference 336
 FT /label= R336I
 FT /note= "wild type Arg replaced with Ile"
 FT Misc-difference 562
 FT /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"

XX WO9740145-A1.

XX PD 30-OCT-1997.

XX PF 24-APR-1997; 97WO-US06563.

XX PR 15-MAY-1996; 960S-0017785.

XX PR 24-APR-1996; 960S-0016117.

XX PA (UNMI) UNIV MICHIGAN.

XX PI Amato K, Kaufman RJ, Pipe SW;

XX DR WPI; 1997-535830/49.

XX PT Modified human pro-coagulant active factor VIII - can be

XX PT administered to haemophiliacs, i.e. factor VIII replacement therapy

XX XX Claim 18; Page -: 57pp; English.

XX The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 CC von Willebrand factor binding site, mutations R336I, R562K and R740A and
 CC an addition of an amino acid sequence spacer between the A2 and A3
 CC domains. Factor VIII, along with calcium and phospholipid, acts as a
 CC cofactor for factor IXA, when it converts factor X to the activated form
 CC (factor XA). FVIII is the coagulation factor deficient in the
 CC X-chromosome-linked bleeding disorder haemophilia A. Several other
 CC mutant FVIII proteins have also been created (see AAW33222-29). The FVIII
 CC mutant F309S (AAW33225) is capable of recombinant secretion at higher
 CC levels than typically obtained with wild type FVIII and retains
 CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
 CC (AAW33223) are resistant to activated protein C (APC) cleavage. The FVIII
 CC mutant comprising a deletion of the B domain and von Willebrand factor
 CC binding site, a mutation at Arg740 and an addition of an amino acid
 CC sequence spacer between the A2 and A3 domains can form a more stable
 CC configuration, and have an approximate 5-fold increase in specific
 CC activity compared to purified wild type FVIII, while increasing their
 CC binding affinity to von Willebrand factor improves their stability.
 CC The FVIII proteins can be administered to haemophiliacs, i.e. FVIII
 CC replacement therapy, while the nucleic acid molecule can be used for
 CC gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

XX SQ Sequence 1383 AA;

Query Match 100.0%; Score 155; DB 18; Length 1383;
 Best Local Similarity 96.7%; Pred. No. 8.4e-14;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
 |||||...|||||...|||||...|||||...
 Db 292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

RESULT 14
 AAW33229
 ID AAW33229 standard; protein; 1383 AA.
 AC AAW33229;
 XX
 DT 30-APR-1998 (first entry)
 XX
 DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
 XX
 KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..346
 FT /note= "factor VIIIA heavy chain"
 FT Region 741..1383
 FT /note= "factor VIIIA light chain"
 FT Domain 1..329
 FT /note= "A1 domain"
 FT Domain 1..179
 FT /note= "plastocyanin-like domain 1"
 FT Domain 187..329
 FT /note= "plastocyanin-like domain 2"
 FT Domain 380..711
 FT /note= "A2 domain"
 FT Misc_feature 711..746
 FT /note= "A spacer of the sequence
 SFQNSRHPSTRQKFNATTPENDIEKTPWF
 AHRTPMPKIQNVSSDLMLL is inserted
 between domains A2 and A3"
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 746..1073
 FT /note= "A3 domain"
 FT Domain 1073..1221
 FT /note= "C1 domain"
 FT Domain 1226..1378
 FT /note= "C2 domain"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Misc-difference 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 XX
 PN W09740145-A1.
 XX
 PD 30-OCT-1997.
 XX
 PF 24-APR-1997; 97WO-US06563.
 XX

PR 15-MAY-1996; 96US-0017785.
 PR 24-APR-1996; 96US-0016117.
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Amano K, Kaufman RJ, Pipe SW;
 DR WPI; 1997-535830/49.
 XX
 XX Modified human pro-coagulant active factor VIII - can be
 PT administered to haemophiliacs, i.e. factor VIII replacement therapy
 XX
 PS Claim 19; Page -: 57pp; English.
 XX
 CC The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 CC von Willebrand factor binding site, mutations F309S, R740A and addition
 CC of an amino acid sequence spacer between the A2 and A3 domains. Factor
 CC VIII, along with calcium and phospholipid, acts as a cofactor for factor
 CC IXA, when it converts factor X to the activated form (factor XA). FVIII
 CC is the coagulation factor deficient in the X-chromosome-linked bleeding
 CC disorder haemophilia A. Several other mutant FVIII proteins have also
 CC been created (see AAW33222-23). The FVIII mutant F309S (AAW33225) is
 CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII
 CC mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated
 CC protein C (APC) cleavage. The present FVIII mutant can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.
 XX
 XX Sequence 1383 AA;
 SQ

Query Match 100.0%; Score 155; DB 18; Length 1383;
 Best Local Similarity 96.7%; Pred. No. 8.4e-14;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
 |||||...|||||...|||||...|||||...
 Db 292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

RESULT 15
 AAP80268
 ID AAP80268 standard; protein; 1424 AA.
 XX
 AC AAP80268;
 XX
 DT 25-MAR-2003 (updated)
 DT 10-OCT-1990 (first entry)
 XX
 DE Modified factor VIII:C sequence with the R740-D1658 deletion.
 XX
 KW Modified factor VIII:C; maturation polypeptide; haemophilia;
 KW blood coagulation; RD deletion; procoagulant.
 XX
 OS Homo sapiens.
 XX
 XX W088000831-A.
 PN
 XX
 PD 11-FEB-1988.
 XX
 PF 31-JUL-1987; 87WO-US01814.
 XX
 XX 01-AUG-1986; 86US-0893375.
 XX
 XX (BIOJ) BIOGEN NV.
 PA
 PA (PASE/) PASEK M P.

```

XX      Pasek MP;
PI
XX
XX      WPT: 1988-049866/07.
DR      N-PSDB; AAN80447.
XX
XX      New DNA sequences encoding modified factor VIII:C - with deletion of DNA
PT      encoding maturation polypeptide, useful for high yield transformation.
XX
XX      Claim 3; Page 60-61-62-63; 97pp; English.
XX
XX      The RD deletion removes the DNA from Ser 741 to Ser 1657.
CC      A major part of the sequence encoding the maturation polypeptide of
CC      factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.
CC      The full length Factor VIII:C cDNA has two changes with respect to the
CC      published sequence (EPO application 160457):
CC      CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
CC      (Phe to Leu). The product is produced in approx. 20 times higher
CC      yields than previous recombinant produced factor VIII:C and are more
CC      easily purified. The peptide is used for treating haemophilia A, both
CC      acute and prolonged bleeding.
CC      See also AAN80444 and AAN80446.
CC      (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      SQ      Sequence      1424 AA;

Query Match      100.0%; Score 155; DB 9; Length 1424;
Best Local Similarity 96.7%; Pred. No. 8.6e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      1      TFLTAQTLLMDLGQFLXCHISSHQDGM 30
          |||||
Db      292  TFLTAQTLLMDLGQFLXCHISSHQDGM 321

```

Search completed: October 21, 2003, 18:53:29
 Job time : 50.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:50:07 ; Search time 16 Seconds
(without alignments)
79.333 Million cell updates/sec

Title: FVIII_PHE328X
Perfect score: 155
Sequence: 1 TFLTAQLMLDGLGQFLXCHISSHQHDGME 30

Scoring table: BLOSUM62X
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Query ⁸		Length	DB	ID	Description
	Score	Match				
1	155	100.0	1438	4	US-09-209-916-1	Sequence 1, Appli
2	155	100.0	1471	1	US-08-683-839B-3	Sequence 3, Appli
3	155	100.0	1661	2	US-08-882-083-2	Sequence 2, Appli
4	155	100.0	1661	2	US-08-558-107-2	Sequence 2, Appli
5	155	100.0	1661	3	US-09-243-539-2	Sequence 2, Appli
6	155	100.0	2332	1	US-07-864-004B-4	Sequence 4, Appli
7	155	100.0	2332	1	US-08-251-937A-4	Sequence 4, Appli
8	155	100.0	2332	1	US-08-212-133A-2	Sequence 2, Appli
9	155	100.0	2332	1	US-08-276-594A-2	Sequence 2, Appli
10	155	100.0	2332	1	US-08-474-503-2	Sequence 2, Appli
11	155	100.0	2332	2	US-08-670-707A-2	Sequence 2, Appli
12	155	100.0	2332	3	US-09-037-601-2	Sequence 2, Appli
13	155	100.0	2332	3	US-09-324-867-3	Sequence 3, Appli
14	155	100.0	2332	4	US-09-315-179-2	Sequence 2, Appli
15	155	100.0	2332	4	US-09-523-656-2	Sequence 2, Appli
16	155	100.0	2332	5	PCT-US93-03275-4	Sequence 4, Appli
17	155	100.0	2332	5	PCT-US94-13200-2	Sequence 2, Appli
18	155	100.0	2351	1	US-08-121-202-2	Sequence 2, Appli
19	155	100.0	2351	1	US-08-366-851A-2	Sequence 2, Appli
20	155	100.0	2351	6	5171844-2	Patent No. 5171844
21	155	100.0	2351	6	5422260-1	Patent No. 5422260
22	147	94.8	2304	3	US-09-324-867-4	Sequence 4, Appli
23	147	94.8	2319	1	US-08-212-133A-8	Sequence 8, Appli
24	147	94.8	2319	1	US-08-474-503-6	Sequence 6, Appli
25	147	94.8	2319	2	US-08-670-707A-6	Sequence 6, Appli
26	147	94.8	2319	3	US-09-037-601-6	Sequence 6, Appli
27	147	94.8	2319	4	US-09-315-179-6	Sequence 6, Appli

Sequence 28, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 39, Appli
Sequence 38, Appli
Sequence 39, Appli
Sequence 38, Appli
Sequence 5, Appli
Sequence 37, Appli
Sequence 37, Appli
Sequence 30, Appli
Sequence 707, App
Sequence 8029, Ap
Sequence 2946, Ap
Sequence 6473, Ap
Sequence 5, Appli

28 147 94.8 2319 4 US-09-523-656-28
29 147 94.8 2319 5 PCT-US94-13200-6
30 146 94.2 2343 3 US-09-324-867-2
31 139 89.7 541 1 US-08-121-202-4
32 139 89.7 1443 2 US-08-670-707A-39
33 139 89.7 1443 3 US-09-037-601-39
34 139 89.7 1443 4 US-09-315-179-39
35 139 89.7 1467 4 US-09-523-656-38
36 139 89.7 2115 3 US-09-324-867-5
37 139 89.7 2133 2 US-08-670-707A-37
38 139 89.7 2133 3 US-09-037-601-37
39 139 89.7 2133 4 US-09-315-179-37
40 139 89.7 2133 4 US-09-523-656-30
41 68 43.9 430 4 US-09-149-476-707
42 54 34.8 635 4 US-09-328-352-8029
43 53 34.2 322 4 US-09-134-001C-2946
44 52 33.5 283 4 US-09-328-352-6473
45 50 32.3 186 1 US-07-960-981-5

ALIGNMENTS

RESULT 1
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

Query Match 100.0%; Score 155; DB 4; Length 1438;
Best Local Similarity 96.7%; Pred. No. 8.8e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQLMLDGLGQFLXCHISSHQHDGME 30
|||||
Db 292 TFLTAQLMLDGLGQFLXCHISSHQHDGME 321

RESULT 2
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: Ili, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of
; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683.839B
FILING DATE: 11-MARCH-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Remillard, Jane E.
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: TTI-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-839B-3

Query Match 100.0%; Score 155; DB 1; Length 1471;
Best Local Similarity 96.7%; Pred. No. 9, 1e-15;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLGGFLXCHISSHQHDGME 30
|||||
DB 311 TFLTAQTLMDLGGFLXCHISSHQHDGME 340

RESULT 3
US-08-882-083-2
Sequence 2, Application US/08882083
Patent No. 5869292
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882.083
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-882-083-2
Query Match 100.0%; Score 155; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 1e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TFLTAQTLMDLGGFLXCHISSHQHDGME 30
|||||
DB 311 TFLTAQTLMDLGGFLXCHISSHQHDGME 340

RESULT 4
US-08-558-107-2
Sequence 2, Application US/08558107
Patent No. 5910481
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,107
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-107-2
Query Match 100.0%; Score 155; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 1e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TFLTAQTLMDLGGFLXCHISSHQHDGME 30
|||||
DB 311 TFLTAQTLMDLGGFLXCHISSHQHDGME 340

RESULT 5
US-09-243-539-2
Sequence 2, Application US/09243539
Patent No. 6130203
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:


```

; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-243-539-2

Query Match 100.0%; Score 155; DB 3; Length 1661;
Best Local Similarity 96.7%; Pred. No. 1e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLXCHSISSHQDGM 30
| | | | | | | | | | | | | | | | | |
Db 311 TELTAQTLLMDLGQFLXCHSISSHQDGM 340

RESULT 6
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EM106
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-07-864-004B-4

Query Match 100.0%; Score 155; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TELTAQTLLMDLGQFLXCHSISSHQDGM 30
| | | | | | | | | | | | | | | | | |
Db 292 TELTAQTLLMDLGQFLXCHSISSHQDGM 321

RESULT 7
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EM106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien

```

```

; TISSUE TYPE: Liver cdna sequence
; US-08-251-937A-4
;
; Query Match          100.0%; Score 155; DB 1; Length 2332;
; Best Local Similarity 96.7%; Pred.No.1.5e-14;
; Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
; US-08-212-133A-2
;
; Query Match          100.0%; Score 155; DB 1; Length 2332;
; Best Local Similarity 96.7%; Pred.No.1.5e-14;
; Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 9
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-594A-2
;
; Query Match          100.0%; Score 155; DB 1; Length 2332;
; Best Local Similarity 96.7%; Pred.No.1.5e-14;
; Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TELTAQTLLMDLGQFLXCHISSHQHDGME 321

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
; US-08-474-503-2

```

Query Match 100.0%; Score 155; DB 1; Length 2332;

Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

```

RESULT 11

```

; US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; US-08-670-707A-2

```

Query Match 100.0%; Score 155; DB 2; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
Db 292 TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

```

RESULT 12

```

; US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; US-09-037-601-2

```

```

Query Match      100.0%; Score 155; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
      |||||||
Db 292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

```

RESULT 13

```

US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:

```

```

; APPLICANT: Lallier, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632ley, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.0010002/JAG/BJD
; CURRENT APPLICATION NUMBER: US/09/324,867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-324-867-3

```

```

Query Match      100.0%; Score 155; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
      |||||||
Db 293 TFLTAQTLLMDLGGQFLXCHSHQHDGME 322

```

RESULT 14

```

US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:

```

```

; APPLICANT: Lollar, John S
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200

```

```

; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-315-179-2

```

```

Query Match      100.0%; Score 155; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
      |||||||
Db 292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

```

RESULT 15

```

US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:

```

```

; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-523-656-2

```

```

Query Match      100.0%; Score 155; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.5e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TFLTAQTLLMDLGGQFLXCHSHQHDGME 30
      |||||||
Db 292 TFLTAQTLLMDLGGQFLXCHSHQHDGME 321

```

```

Search completed: October 21, 2003, 18:59:47
Job time : 17 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:53:43 ; Search time 21.25 Seconds
(without alignments)
236.415 Million cell updates/sec

Title: FVIII-PHE328X

Perfect score: 155

Sequence: 1 TFLTAQTLLMDLQGLLXCHTSSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155	100.0	1438	14	US-10-006-091-1
2	155	100.0	1438	14	US-10-047-257-1
3	155	100.0	1438	15	US-10-225-900-1
4	155	100.0	1471	14	US-10-095-718-2
5	155	100.0	2332	10	US-09-957-641-2
6	155	100.0	2332	12	US-10-131-510A-2
7	155	100.0	2332	15	US-10-187-319-2
8	155	100.0	2351	12	US-10-133-907-4
9	155	100.0	2351	15	US-10-132-829-4
10	155	100.0	2351	15	US-10-172-712-27
11	147	94.8	2319	12	US-10-131-510A-6
12	147	94.8	2319	15	US-10-187-319-6
13	146	94.2	1431	14	US-10-095-718-4
14	139	89.7	1443	12	US-10-131-510A-39
15	139	89.7	1443	15	US-10-187-319-39

16	89.7	2133	12	US-10-131-510A-37	Sequence 37, Appl
17	89.7	2133	15	US-10-187-319-37	Sequence 37, Appl
18	46.5	1160	12	US-10-137-870-234	Sequence 234, App
19	46.5	1160	12	US-10-140-018-234	Sequence 234, App
20	46.5	1160	12	US-10-140-021-234	Sequence 234, App
21	46.5	1160	12	US-10-140-274-234	Sequence 234, App
22	46.5	1160	12	US-10-140-471-234	Sequence 234, App
23	46.5	1160	12	US-10-140-807-234	Sequence 234, App
24	46.5	1160	12	US-10-140-922-234	Sequence 234, App
25	46.5	1160	12	US-10-140-924-234	Sequence 234, App
26	46.5	1160	12	US-10-140-926-234	Sequence 234, App
27	46.5	1160	12	US-10-141-698-234	Sequence 234, App
28	46.5	1160	12	US-10-141-702-234	Sequence 234, App
29	46.5	1160	12	US-10-141-704-234	Sequence 234, App
30	46.5	1160	12	US-10-142-421-234	Sequence 234, App
31	46.5	1160	12	US-10-142-432-234	Sequence 234, App
32	46.5	1160	12	US-10-142-767-234	Sequence 234, App
33	46.5	1160	12	US-10-143-033-234	Sequence 234, App
34	46.5	1160	12	US-10-144-994-234	Sequence 234, App
35	46.5	1160	12	US-10-145-628-234	Sequence 234, App
36	46.5	1160	12	US-10-145-631-234	Sequence 234, App
37	46.5	1160	12	US-10-145-633-234	Sequence 234, App
38	46.5	1160	12	US-10-145-746-234	Sequence 234, App
39	46.5	1160	12	US-10-145-748-234	Sequence 234, App
40	46.5	1160	12	US-10-145-823-234	Sequence 234, App
41	46.5	1160	12	US-10-145-826-234	Sequence 234, App
42	46.5	1160	12	US-10-145-870-234	Sequence 234, App
43	46.5	1160	12	US-10-145-876-234	Sequence 234, App
44	46.5	1160	12	US-10-145-959-234	Sequence 234, App
45	46.5	1160	12	US-10-146-724-234	Sequence 234, App

ALIGNMENTS

RESULT 1

US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match 100.0%; Score 155; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2,1e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQGLLXCHTSSHQHDGME 30
|||||
DB 292 TFLTAQTLLMDLQGLLXCHTSSHQHDGME 321

RESULT 2

US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:

```

; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-10-047-257-1

Query Match          100.0%; Score 155; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2.1e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 321

RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US2003007752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-10-225-900-1

Query Match          100.0%; Score 155; DB 15; Length 1438;
Best Local Similarity 96.7%; Pred. No. 2.1e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 321

RESULT 4
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony

```

```

; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

Query Match          100.0%; Score 155; DB 14; Length 1471;
Best Local Similarity 96.7%; Pred. No. 2.2e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 30
|||||
Db 311 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 340

RESULT 5
US-09-957-641-2
; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00
; CURRENT APPLICATION NUMBER: US/09/957,641
; CURRENT FILING DATE: 2001-09-16
; PRIOR APPLICATION NUMBER: US 60/234047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-957-641-2

Query Match          100.0%; Score 155; DB 10; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.6e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 30
|||||
Db 292 TFLTAQTLLMDLGGQFLXCHISSHQHDGME 321

RESULT 6
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601

```

```

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-510A-2

```

```

Query Match      100.0%; Score 155; DB 12; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.6e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
          |||||||||||||||:|||||||
Db      292  TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

```

```

RESULT 7
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: <Unknown>
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

```

```

Query Match      100.0%; Score 155; DB 15; Length 2332;
Best Local Similarity 96.7%; Pred. No. 3.6e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
          |||||||||||||||:|||||||
Db      292  TFLTAQTLLMDLGQFLXCHISSHQHDGME 321

```

```

RESULT 8
US-10-133-907-4
; Sequence 4, Application US/10133907
; Publication No. US20030195223A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VIII
; FILE REFERENCE: 6627-PALL70
; CURRENT APPLICATION NUMBER: US/10/133,907
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-907-4

```

```

Query Match      100.0%; Score 155; DB 12; Length 2351;
Best Local Similarity 96.7%; Pred. No. 3.7e-13;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  TFLTAQTLLMDLGQFLXCHISSHQHDGME 30
          |||||||||||||||:|||||||
Db      311  TFLTAQTLLMDLGQFLXCHISSHQHDGME 340

```

```

RESULT 9
US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VIII
; FILE REFERENCE: 6627-PALL70
; CURRENT APPLICATION NUMBER: US/10/132,829
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-4

```

Query Match 100.0%; Score 155; DB 15; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 3.7e-13;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLXCHSHQHDGME 30
 |||||
 Db 311 TFLTAQTLMDLQGLXCHSHQHDGME 340

RESULT 10

US-10-172-712-27
 ; Sequence 27, Application US/10172712
 ; Publication No. US20030125232A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRIFFIN, JOHN H.
 ; APPLICANT: GALE, ANDREW J.
 ; APPLICANT: GETZOFF, ELIZABETH D.
 ; APPLICANT: PELLEGUER, JEAN-LUC
 ; TITLE OF INVENTION: STABILIZED PROTEINS WITH ENGINEERED DISULFIDE BONDS
 ; FILE REFERENCE: 4198-4001US1
 ; CURRENT APPLICATION NUMBER: US/10/172,712
 ; CURRENT FILING DATE: 2002-09-30
 ; PRIOR APPLICATION NUMBER: 60/298,578
 ; PRIOR FILING DATE: 2001-06-14
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 27
 ; LENGTH: 2351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-172-712-27

Query Match 100.0%; Score 155; DB 15; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 3.7e-13;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLXCHSHQHDGME 30
 |||||
 Db 311 TFLTAQTLMDLQGLXCHSHQHDGME 340

RESULT 11

US-10-131-510A-6
 ; Sequence 6, Application US/10131510A
 ; Publication No. US20030166536A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LOLLAR, JOHN S.
 ; TITLE OF INVENTION: Modified Factor VIII
 ; FILE REFERENCE: 75-95J
 ; CURRENT APPLICATION NUMBER: US/10/131,510A
 ; CURRENT FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: U.S. 09/315,179
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: U.S. 09/037,601
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: U.S. 08/670,707
 ; PRIOR FILING DATE: 1996-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US97/11155
 ; PRIOR FILING DATE: 1997-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/13200
 ; PRIOR FILING DATE: 1994-11-15
 ; PRIOR APPLICATION NUMBER: U.S. 08/212,133
 ; PRIOR FILING DATE: 1994-03-11
 ; PRIOR APPLICATION NUMBER: U.S. 07/864,004
 ; PRIOR FILING DATE: 1992-04-07
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 2319
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-131-510A-6

Query Match 94.8%; Score 147; DB 12; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 5e-12;
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLXCHSHQHDGME 30
 |||||
 Db 312 TFLTAQTLMDLQGLXCHSHQHDGME 341

RESULT 12

US-10-187-319-6
 ; GENERAL INFORMATION:
 ; APPLICANT: LOLLAR, JOHN S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA: US/10/187,319
 ; APPLICATION NUMBER: US/10/187,319
 ; FILING DATE: 27-Aug-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/523,656
 ; FILING DATE: 2000-03-10
 ; APPLICATION NUMBER: US 09/037,601
 ; FILING DATE: 1998-03-10
 ; APPLICATION NUMBER: WO PCT/US97/11155
 ; FILING DATE: 1997-06-26
 ; APPLICATION NUMBER: US 08/670,707
 ; FILING DATE: 1996-06-26
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Greenlee, Lorraine L.
 ; REGISTRATION NUMBER: 27,894
 ; REFERENCE/DOCKET NUMBER: 75-95K
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/499-8080
 ; TELEFAX: 303/499-8089
 ; TITLE: Sequence of the Murine Factor VIII cDNA
 ; JOURNAL: Genomics
 ; VOLUME: 16
 ; PAGES: 374-379
 ; DATE: 1993
 ; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-10-187-319-6

Query Match 94.8%; Score 147; DB 15; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 5e-12;
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLMDLQGLXCHSHQHDGME 30
 |||||
 Db 312 TFLTAQTLMDLQGLXCHSHQHDGME 341

RESULT 13

US-10-095-718-4
 ; Sequence 4, Application US/10095718
 ; Publication No. US20020131956A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walsh, Christopher
 ; APPLICANT: Chao, Hengjun
 ; APPLICANT: Burstein, Haim


```
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: canine B-domain deleted factor VIII
US-10-095-718-4

Query Match      94.2%; Score 146; DB 14; Length 1431;
Best Local Similarity 90.0%; Pred. No. 4.le-12;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 TFLTAQTLLMDLGQFLXCHISSHHQDGM 30
        ||||| ||||| ||||| ||||| |||||
DB      306 TFLTAQTLLMDLGQFLXCHIPSHQDGM 335

RESULT 14
US-10-131-510A-39
; Sequence 39, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: factor VIII lacking
; OTHER INFORMATION: the B domain
US-10-131-510A-39

Query Match      89.7%; Score 139; DB 12; Length 1443;
Best Local Similarity 86.7%; Pred. No. 4.le-11;
Matches 26; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 TFLTAQTLLMDLGQFLXCHISSHHQDGM 30
        ||||| ||||| ||||| ||||| |||||
DB      312 TFLTAQTLLMDLGQFLXCHISSHHQDGM 341
```

```
Search completed: October 21, 2003, 19:01:20
Job time : 23.25 secs
```

```
RESULT 15
US-10-187-319-39
; Sequence 39, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-10-187-319-39

Query Match      89.7%; Score 139; DB 15; Length 1443;
Best Local Similarity 86.7%; Pred. No. 4.le-11;
Matches 26; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 TFLTAQTLLMDLGQFLXCHISSHHQDGM 30
        ||||| ||||| ||||| ||||| |||||
DB      312 TFLTAQTLLMDLGQFLXCHISSHHQDGM 341

Search completed: October 21, 2003, 19:01:20
Job time : 23.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:49:52 ; Search time 17.5 seconds
(without alignments)
164.861 Million cell updates/sec

Title: FVIII_PHE328X

Perfect score: 155

Sequence: 1 TFLTAQTLLMDLGLQLXCHISSHQHDGME 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155	100.0	2351	1 EZHU	coagulation factor
2	147	94.8	2131	2 A47004	coagulation factor
3	139	89.7	2133	2 T42763	coagulation factor
4	64.5	41.6	1069	1 KRUH	ferroxidase (EC 1.10.3)
5	62	40.0	1059	1 A35210	L-ascorbate oxidase (EC 1.10.3)
6	61	39.4	721	2 H82528	laccase (EC 1.10.3)
7	54	34.8	609	1 KSASL1	seryl-tRNA synthetase
8	52	33.5	427	2 A12785	periplasmic cell d
9	52	33.5	435	2 C97565	myc2 protein - rat
10	51	32.9	527	2 C70397	oncoprotein zc-Myc
11	50	32.3	398	2 S99005	transforming prote
12	50	32.3	406	2 A48059	transforming prote
13	50	32.3	414	1 TVTRMC	transforming prote
14	50	32.3	438	2 JG1179	transforming prote
15	50	32.3	439	1 TVHUM	transforming prote
16	50	32.3	439	1 TVHMS	transforming prote
17	50	32.3	439	1 TVRMC	transforming prote
18	50	32.3	439	2 JG1178	transforming prote
19	50	32.3	439	2 JU0449	c-myc protein - ch
20	50	32.3	440	4 TVHUT	transforming prote
21	50	32.3	465	2 S03325	transforming prote
22	49.5	31.9	137	2 T07891	protein kinase (BC
23	49	31.6	379	2 F82963	hypothetical prote
24	49	31.6	379	2 AH0542	conserved hypothe
25	48	31.0	487	2 D71423	hypothetical prote
26	48	31.0	519	2 H97724	multidrug resistan
27	48	31.0	554	2 T51213	hypothetical prote
28	48	31.0	567	2 T44928	L-ascorbate oxidas
29	47.5	30.6	213	2 S29910	hypothetical prote

30	47.5	30.6	230	2 T37412	probable 27.1k pro
31	47.5	30.6	260	2 H34284	cytochrome-c oxida
32	47.5	30.6	260	2 T11798	cytochrome-c oxida
33	47.5	30.6	304	2 S76572	hypothetical prote
34	47.5	30.6	322	1 WNV210	A-type inclusion p
35	47	30.3	142	2 G95329	probable dehydroge
36	47	30.3	143	2 H64000	hypothetical prote
37	47	30.3	311	2 A95342	hypothetical prote
38	47	30.3	343	2 T35030	probable copper ox
39	47	30.3	552	2 A51027	L-ascorbate oxidas
40	47	30.3	569	2 E84904	probable laccase (
41	47	30.3	578	2 S66353	L-ascorbate oxidas
42	47	30.3	579	2 S11027	L-ascorbate oxidas
43	47	30.3	587	1 KSKVAO	hypothetical prote
44	47	30.3	608	2 S69564	probable multicopp
45	47	30.3	622	2 S62580	

ALIGNMENTS

RESULT 1

EZHU

coagulation factor VIII precursor [validated] - human
N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000
C:Accession: I54318; A00525; I58059; A23584; A26174; A4348; A43986; S63527; S66445
R:Gitschler, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.
A:Reference number: I54318; MUID:93265012; PMID:1303178
A:Accession: I54318
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1921, 'S', 1923-2351 <RES>
A:Cross-references: GB:M88648; NID:gl82381; PIDN:AAAS2420.1; PID:gl82383
R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Se
Nature 312, 330-337, 1984
A:Title: Expression of active human factor VIII from recombinant DNA clones.
A:Reference number: A00525; MUID:85061548; PMID:6438526
A:Accession: A00525
A:Molecule type: mRNA
A:Residues: 1-2351 <WOO>
A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R:Footie, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, I
S, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A:Reference number: I58059; MUID:85061550; PMID:6438528
A:Accession: I58059
A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
A:Cross-references: GB:M14113; NID:gl82802; PIDN:AAAS2484.1; PID:gl82803
R:Triuett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.;
B.; Randolph, A.; Ordeu, M.S.; Valenzuela, P.; Dahl, H.H.; Favallaro, J.; Hansen, C
DNA 4, 333-349, 1985
A:Title: Characterization of the polypeptide composition of human factor VIII:C and
A:Reference number: A23584; MUID:86081164; PMID:3935400
A:Accession: A23584
A:Molecule type: mRNA
A:Residues: 1-2351 <TRU>
A:Cross-references: GB:M14113; NID:gl82817; PIDN:AAAS2485.1; PID:gl82818
R:Eaton, D.; Rodriguez, H.; Vohar, G.A.
Biochemistry 25, 505-512, 1986
A:Title: Proteolytic processing of human factor VIII. Correlation of specific clea
ity.
A:Reference number: A26174; MUID:86159740; PMID:3082357
A:Accession: A26174
A:Molecule type: protein
A:Residues: 20-36;392-399, 'X', 401-402;1668-1678;1709-1722, 'D', 1723-1725;1741-1755 <
R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
 A:Title: Identification and functional importance of tyrosine sulfate residues within re
 A:Reference number: A42348; MUID:92207952; PMID:1554716
 A:Accession: A42348
 A:Molecule type: protein
 A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
 A:Experimental source: recombinant material from Chinese hamster ovary cells
 A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 R:Pay, P.J.; Smudzin, J.M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A:Reference number: A43986; MUID:89340500; PMID:2503509
 A:Accession: A43986
 A:Molecule type: protein
 A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>
 R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
 J. Biol. Chem. 266, 740-746, 1991
 A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A:Reference number: A56109; MUID:91093266; PMID:1898735
 A:Accession: A56109
 A:Contents: annotation; sulfation
 R:Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
 Nature 312, 326-330, 1984
 A:Title: Characterization of the human factor VIII gene.
 A:Reference number: A56196; MUID:85061547; PMID:6438525
 A:Accession: A56196
 A:Contents: annotation; introns
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o
 A:Reference number: A56216; MUID:95338127; PMID:7613471
 A:Accession: A56216
 A:Contents: annotation; disulfide bonds
 A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
 R:Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A:Reference number: S63527; MUID:96163459; PMID:8575434
 A:Accession: S63527
 A:Molecule type: protein
 A:Residues: 733-752;753-759 <KJA>
 R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baackman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A:Reference number: S66445; MUID:96048024; PMID:7556150
 A:Accession: S66445
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1668-1685 <LJN>
 C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
 C:Genetics:
 A:Gene: GDB:F8C
 A:Cross-references: GDB:119124; OMIM:306700
 A:Map position: Xq28-Xq28
 A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C:Function:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 C:Pathway: blood coagulation
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
 F:20-740/Product: coagulation factor VIIa heavy chain #status experimental <ACH>
 F:20-356/Domain: A1 <DAL>
 F:23-348/Domain: ferroxidase repeat homology <FOI>
 F:392-759/Domain: A2 <DA2>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:760-1667/Domain: B <DBO>
 F:1668-2351/Product: coagulation factor VIIa light chain #status experimental <ACL>
 F:1709-2038/Domain: A3 <DA3>
 F:1716-2038/Domain: ferroxidase repeat homology <FO3>
 F:2039-2191/Domain: C1 <DC1>
 F:2039-2188/Domain: discoidin I amino-terminal homology <DNI>
 F:2192-2351/Domain: C2 <DC2>
 F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>
 F:60,258,601,776,803,847,913,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #s
 F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predict
 F:365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experime
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experi
 F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
 F:2193-2345/Disulfide bonds: #status predicted
 Query Match 100.0%; Score 155; DB 1; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 3e-14; Mismatches 0; Indels 0; Gaps 0;
 Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TFLTATQLTMDLQGLLXCHSHSHQHDGME 30
 Db 311 TELTAQTLLMDLQGLLXCHSHSHQHDGME 340
 RESULT 2
 A47004
 coagulation factor VIII precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 C:Accession: A47004
 R:Elder, B.; Lakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993
 A:Title: Sequence of the murine factor VIII cDNA.
 A:Reference number: A47004; MUID:93300511; PMID:8314577
 A:Accession: A47004
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2319 <ELD>
 A:Cross-references: GB:L05573; NID:gl92456; PIDN:AAA37385.1; PID:gl92457
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferrox
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:23-349/Domain: ferroxidase repeat homology <FOI>
 F:402-730/Domain: ferroxidase repeat homology <FO2>
 F:1686-2006/Domain: ferroxidase repeat homology <FO3>
 F:2007-2156/Domain: discoidin I amino-terminal homology <DNI>
 F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>
 Query Match 94.8%; Score 147; DB 2; Length 2319;
 Best Local Similarity 90.0%; Pred. No. 4.5e-13; Mismatches 0; Indels 0; Gaps 0;
 Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TFLTATQLTMDLQGLLXCHSHSHQHDGME 30
 Db 312 TELTAQTLLMDLQGLLXCHSHSHQHDGME 341
 RESULT 3
 T42763
 coagulation factor VIII precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42763
 R:Lollar, P.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z22269
 A:Accession: T42763
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2133 <LOL>
 A:Cross-references: EMBL:U49517; NID:gl511633; PID:gl511634; PIDN:AA06705.1
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferrox
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
 F:23-349/Domain: ferroxidase repeat homology <FOX1>
 F:402-730/Domain: ferroxidase repeat homology <FOX2>
 F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

1

A:Gene: c-myc

A:Introns: 253/1

C:Superfamily: myc transforming protein; myc transforming protein homology

C:Keywords: DNA binding; transforming protein

F:14-436/Domain: myc transforming protein homology <MYC>

Query Match

Best Local Similarity 32.3%; Score 50; DB 2; Length 438;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 16 LLXCHISSHQHD 27
I:|||||:

Db 296 LKRCVSTQHN 307

RESULT 15

TVHUM

Transforming protein C-myc - human

N:Alternate names: p64 myc oncogene

C:Species: Homo sapiens (man)

C>Date: 14-Nov-1983 #sequence_revision 31-Mar-1992 #text_change 24-Nov-1999

C:Accession: A01349; A26245; A26246; I67315; I53224; I61947; I38052; I57605; I37

R:Watson, D.K.; Psailidopoulos, M.C.; Samuel, K.P.; Dalla-Pavera, R.; Papas, T.S.

Proc. Natl. Acad. Sci. U.S.A. 80, 3642-3645, 1983

A:Title: Nucleotide sequence analysis of human c-myc locus, chicken homologue, and myelo

A:Reference number: A01349; MUID:83221652; PMID:6304729

A:Accession: A01349

A:Molecule type: DNA

A:Residues: 1-439 <WATS>

A:Cross-references: GB:K01906; NID:g188937; PIDN:AAA59881.1; PID:g386968

R:Colby, W.W.; Chen, E.Y.; Smith, D.H.; Levinson, A.D.

Nature 301, 722-725, 1983

A:Title: Identification and nucleotide sequence of a human locus homologous to the v-myc

A:Reference number: A26245; MUID:83141777; PMID:6298632

A:Accession: A26245

A:Molecule type: DNA

A:Residues: 1-439 <COL>

A:Cross-references: GB:J00120; GB:K01908; GB:M23541; GB:V00501; GB:X00364; NID:g515632;

R:Watt, R.; Stanton, L.W.; Marcu, K.B.; Gallo, R.C.; Croce, C.M.; Rovera, G.

Nature 303, 725-728, 1983

A:Title: Nucleotide sequence of cloned cDNA of human c-myc oncogene.

A:Reference number: A26246; MUID:83219310; PMID:6304538

A:Accession: A26246

A:Molecule type: mRNA

A:Residues: 1-439 <WATT>

A:Cross-references: GB:V00568; NID:g34815; PIDN:CAA23831.1; PID:g34816

R:Battery, J.; Moulding, C.; Taub, R.; Murphy, W.; Stewart, T.; Potter, H.; Lenoir, G.; I

Cell 34, 779-787, 1983

A:Title: The human c-myc oncogene: structural consequences of translocation into the Igh

A:Reference number: A21061; MUID:84026482; PMID:6414718

A:Accession: A21061

A:Molecule type: DNA

A:Residues: 1-439 <BAT>

A:Cross-references: NID:g188917; PIDN:AAA59880.1; PID:g188919

R:Gazin, C.; De Dinechin, S.D.; Hampe, A.; Masson, J.

EMBO J. 3, 383-387, 1984

A:Title: Nucleotide sequence of the human c-myc locus: provocative open reading frame w

A:Reference number: I53224; MUID:84182501; PMID:6714223

A:Accession: I67315

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-439 <GAT>

A:Cross-references: NID:g188917; PIDN:AAA59880.1; PID:g188919

R:Gazin, C.; De Dinechin, S.D.; Hampe, A.; Masson, J.

EMBO J. 3, 383-387, 1984

A:Title: Nucleotide sequence of the human c-myc locus: provocative open reading frame w

A:Reference number: I53224; MUID:84182501; PMID:6714223

A:Accession: I67315

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-439 <GAT>

A:Cross-references: GB:D10493; NID:g219932; PIDN:BAA01375.1; PID:g219934

A:Accession: I53224

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-439 <GAT>

A:Cross-references: GB:D10493; NID:g219932; PIDN:BAA01375.1; PID:g219934

R:Bernard, O.; Cory, S.; Gerondakis, S.; Webb, E.; Adams, J.M.

EMBO J. 2, 2375-2383, 1983

A:Title: Sequence of the murine and human cellular myc oncogenes and two modes of myc tr

A:Reference number: I38052; MUID:84131953; PMID:6321164

A:Accession: I61947

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-439 <BER1>

A:Cross-references: GB:K01904; NID:g188931; PIDN:AAA36341.1; PID:g386967

A:Accession: I38052

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-252 <BER2>

A:Cross-references: EMBL:X00196; NID:g34822; PIDN:CAA25015.1; PID:g34823

R:Bentley, D.L.; Groudine, M.

Mol. Cell. Biol. 6, 3481-3489, 1986

A:Title: Novel promoter upstream of the human c-myc gene and regulation of c-myc ex

A:Reference number: I57605; MUID:87089682; PMID:3540591

A:Accession: I57605

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-170 <BEN>

A:Cross-references: GB:M13929; NID:g188963; PIDN:AAA88092.1; PID:g188964

R:Rabbits, T.H.; Forster, A.; Hamlyn, P.; Baer, R.

Nature 309, 592-597, 1984

A:Title: Effect of somatic mutation within translocated c-myc genes in Burkitt's ly

A:Reference number: I37244; MUID:84219769; PMID:6547209

A:Accession: I37244

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-61, 'p', 63-252 <RAB>

A:Cross-references: EMBL:X00676; NID:g29996; PIDN:CAA25288.1; PID:g29997

R:Hayday, A.C.; Gillies, S.D.; Saito, H.; Wood, C.; Wiman, K.; Hayward, W.S.; Toney

Nature 307, 334-340, 1984

A:Title: Activation of a translocated human c-myc gene by an enhancer in the immuno

A:Reference number: I38075; MUID:84117466; PMID:6420706

A:Accession: I38075

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-7 <HAY>

A:Cross-references: EMBL:X00344; NID:g35146; PIDN:CAA25095.1; PID:g580340

R:Hollis, G.F.; Gazdar, A.F.; Bertness, V.L.; Kirsch, I.R.

Mol. Cell. Biol. 8, 124-129, 1988

A:Title: Complex translocation disrupts c-myc regulation in a human plasma cell mye

A:Reference number: I57633; MUID:88094377; PMID:3275865

A:Accession: I57633

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 437-439 <HOL>

A:Cross-references: GB:M19724; NID:g190971

A:Note: this translation is not annotated in GenBank entry HUMRCMYC3, release 112.0

C:Genetics:

A:Gene: GDB:MYC

A:Cross-references: GDB:120208; OMIM:190080

A:Map position: 8q24.12-8q24.13

A:Introns: 253/1

C:Superfamily: myc transforming protein; myc transforming protein homology

C:Keywords: DNA binding; glycoprotein; leucine zipper; transcription regulation; tr

F:14-437/Domain: myc transforming protein homology <MYC>

F:406-434/Region: leucine zipper motif

F:58/Binding site: carbohydrate (Thr) (covalent) #status predicted

Query Match

Best Local Similarity 32.3%; Score 50; DB 1; Length 439;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 16 LLXCHISSHQHD 27
I:|||||:

Db 297 LKRCVSTQHN 308

Search completed: October 21, 2003, 18:58:37

Job time: 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:47:22 ; Search time 9.25 Seconds
(without alignments)
152.519 Million cell updates/sec

Title: FVIII_PHE328X

Perfect score: 155

Sequence: 1 TFLTAQTLLMDLGGFLXCHTSSHQHGMG 30

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	155	100.0	2351	1 F88_HUMAN	P00451 homo sapien
2	147	94.8	2319	1 F88_MOUSE	Q06194 mus musculu
3	139	89.7	2133	1 F88_PIG	P12263 sus scrofa
4	64.5	41.6	1065	1 CERU_HUMAN	P00450 homo sapien
5	63	40.6	1062	1 CERU_MOUSE	Q61147 mus musculu
6	62	40.0	1059	1 CERU_RAT	P13635 ratus norv
7	54	34.8	609	1 LACL_EMENI	P17489 emericella
8	52	33.5	427	1 SYS_AGR5	Q8ueq2 agrobacteri
9	50	32.3	394	1 MYC1_CYPCA	P490341 cyprinus ca
10	50	32.3	399	1 MYC2_CYPCA	P49709 carassius a
11	50	32.3	401	1 MYC2_CYPCA	Q90342 cyprinus ca
12	50	32.3	406	1 MYC2_CYPCA	P52160 brachydanio
13	50	32.3	414	1 MYC2_CYPCA	P06646 oncorhynch
14	50	32.3	438	1 MYC2_CYPCA	P49032 callithrix
15	50	32.3	439	1 MYC2_CYPCA	Q28350 canis fami
16	50	32.3	439	1 MYC2_CYPCA	P01106 homo sapien
17	50	32.3	439	1 MYC2_CYPCA	P49033 hyllobates l
18	50	32.3	439	1 MYC2_CYPCA	P22555 marmota non
19	50	32.3	439	1 MYC2_CYPCA	P01108 mus musculu
20	50	32.3	439	1 MYC2_CYPCA	P23583 pan troglod
21	50	32.3	439	1 MYC2_CYPCA	Q29031 sus scrofa
22	50	32.3	439	1 MYC2_CYPCA	P09416 ratus norv
23	50	32.3	439	1 MYC2_CYPCA	Q28566 ovis aries
24	49	31.6	642	1 PHSA_STRAT	Q53692 streptomyce
25	48	31.0	486	1 LACL_BOTCI	Q12570 botrytis ci
26	47.5	30.6	109	1 COX3_ARBLI	P24758 vaccinia lix
27	47.5	30.6	213	1 AT12_VACCV	P12702 paracentrot
28	47.5	30.6	261	1 COX3_PARLI	P54004 synecocyst
29	47.5	30.6	304	1 Y199_SYNY3	P21114 vaccinia vi
30	47.5	30.6	322	1 AT12_VACCC	P43936 haemophilus
31	47	30.3	143	1 Y1FN_HAEIN	Q9W758 xenopus lae
32	47	30.3	437	1 C1BQ_XENLA	P37064 cucurbita p
33	47	30.3	552	1 ASO_CUCPM	

34	47	30.3	578	1 ASO_TOBAC	Q40588 nicotiana t
35	47	30.3	579	1 ASO_CUCMA	P24792 cucurbita m
36	47	30.3	587	1 ASO_CUCSA	P14133 cucumis sat
37	47	30.3	608	1 YD56_YEAST	Q04399 saccharomyc
38	47	30.3	622	1 YAK8_SCHPO	Q09320 schizosacch
39	47	30.3	839	1 V2A_CWVQ	P06012 cucurbit mo
40	46.5	30.0	205	1 FLRE_HUMAN	P30043 homo sapien
41	46.5	30.0	266	1 TRPC_METJA	Q58328 methanococc
42	46.5	30.0	281	1 COX3_RHIST	P80441 rhizopus st
43	46.5	30.0	797	1 T4AP_HUMAN	Q8tel6 homo sapien
44	46	29.7	474	1 T1L1_MOUSE	Q923u0 mus musculu
45	46	29.7	513	1 THD1_PASMO	Q9ckj2 pasteurella

ALIGNMENTS

RESULT 1
F88_HUMAN STANDARD; PRT; 2351 AA.
AC P00451;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component)
DE (Antithemophilic factor) (AHF).
GN F8 OR F8C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Masiaz F.R., Merryweather J.P., Najarian R.,
RA Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Urdea M.S., Valenzuela P., Dahl H.-H.N., Favalaro J., Hansen J.,
RA Nordfang O., Ezban M.;
RA "Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA";
RL DNA 4:333-349(1985).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,
RA Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RL "Expression of active human factor VIII from recombinant DNA clones.";
RL Nature 312:330-337(1984).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA Hewick R.M.;
RT "Molecular cloning of a cDNA encoding human antihaemophilic factor.";
RL Nature 312:342-347(1984).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93265012; PubMed=1303178;
RA Gitschier J., Wood W.I.;
RT "Sequence of the exon-containing regions of the human factor VIII
RT gene.";
RN Hum. Mol. Genet. 1:199-200(1992).
[5]
RP SEQUENCE OF 2064-2070 FROM N.A.
RA de Water N.S., Williams R., Browett P.J.;
RL submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[6]
RP SULFATION OF TYR-1699.

RX MEDLINE=91093266; PubMed=1898735;
RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeest M.P.,
RA Mertens K., van Mourik J.A.;
RT "Sulfation of Tyr-680 of human blood coagulation factor VIII is
RT essential for the interaction of factor VIII with von Willebrand
RT factor";
RL J. Biol. Chem. 266:740-746(1991).
RN [7]
RN SULTATION.
RX MEDLINE=92207952; PubMed=1554716;
RA Pittman D.D., Wang J.H., Kaufman R.J.;
RT "Identification and functional importance of tyrosine sulfate
RT residues within recombinant factor VIII";
RL Biochemistry 31:3315-3325(1992).
RN [8]
RN STRUCTURE BY NMR OF 2322-2343.
RX MEDLINE=95200924; PubMed=7893714;
RA Gilbert G.E., Baleja J.D.;
RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT amphipathic structure as determined by NMR spectroscopy";
RL Biochemistry 34:3022-3031(1995).
RN [9]
RN REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=91221499; PubMed=1902642;
RA Gitschier J.;
RT "The molecular basis of hemophilia A";
RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
RN [10]
RN REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=89088506; PubMed=2491949;
RA White G.C. II, Shoemaker C.B.;
RT "Factor VIII gene and hemophilia A";
RL Blood 73:1-12(1989).
RN [11]
RN REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=95245332; PubMed=7728145;
RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
RT "Molecular etiology of factor VIII deficiency in hemophilia A";
RL Hum. Mutat. 5:1-22(1995).
RN [12]
RN VARIANT HEMA GLN-2326.
RX MEDLINE=86235434; PubMed=3012775;
RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
RT "Identification of a missense mutation in the factor VIII gene of a
RT mild hemophiliac";
RL Science 232:1415-1416(1986).
RN [13]
RN VARIANT HEMA PRO-2135.
RX MEDLINE=88096539; PubMed=3122181;
RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
RT "A novel missense mutation in the factor VIII gene identified by
RT analysis of amplified hemophilia DNA sequences";
RL Nucleic Acids Res. 15:9797-9805(1987).
RN [14]
RN VARIANT HEMA GLN-2228.
RX MEDLINE=88191889; PubMed=2833855;
RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
RA Kazazian H.H.;
RT "Nonsense and missense mutations in hemophilia A: estimate of the
RT relative mutation rate at CG dinucleotides";
RL Am. J. Hum. Genet. 42:718-725(1988).
RN [15]
RN VARIANT HEMA GLY-291.
RX MEDLINE=88220354; PubMed=2835904;
RA Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
RA Antonarakis S.E.;
RT "Moderately severe hemophilia A resulting from Glu->Gly substitution
RT in exon 7 of the factor VIII gene";
RL Am. J. Hum. Genet. 42:867-871(1988).
RN [16]
RN VARIANT HEMA CYS-1708.
RX MEDLINE=89274393; PubMed=2499363;
RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1,689-Cys: a
RT nonfunctional cofactor occurring in a patient with severe hemophilia
RT A";
RL Blood 73:2117-2122(1989).
RN [17]
RN VARIANT HEMA CYS-391.
RX MEDLINE=90001543; PubMed=2506948;
RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
RT "An arginine to cysteine amino acid substitution at a critical
RT thrombin cleavage site in a dysfunctional factor VIII molecule";
RL Blood 74:1612-1617(1989).
RN [18]
RN VARIANT HEMA LEU-189.
RX MEDLINE=90057680; PubMed=2510835;
RA Chan V., Chan T.K., Tong T.M., Todd D.;
RT "A novel missense mutation in exon 4 of the factor VIII:C gene
RT resulting in moderately severe hemophilia A";
RL Blood 74:2688-2691(1989).
RN [19]
RN VARIANT HEMA LEU-2326.
RX MEDLINE=89197216; PubMed=2495245;
RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
RT of the factor VIII gene";
RL Hum. Genet. 81:335-338(1989).
RN [20]
RN VARIANT HEMA HIS-391.
RX MEDLINE=89264602; PubMed=2498882;
RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA Fujimaki M., Hoyer L.W.;
RT "Direct characterization of factor VIII in plasma: detection of a
RT mutation altering a thrombin cleavage site
RT (arginine-372->histidine)";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
RN [21]
RN VARIANT HEMA CYS-1708.
RX MEDLINE=90105723; PubMed=2104766;
RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
RT Cys) in the factor VIII gene of two unrelated patients with
RT cross-reacting material-positive hemophilia A";
RL Blood 75:384-389(1990).
RN [22]
RN VARIANTS HEMA GLN-2228 AND LEU-2326.
RX MEDLINE=90123183; PubMed=2105106;
RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
RT "Recurrent mutations and three novel rearrangements in the factor
RT VIII gene of hemophilia A patients of Italian descent";
RL Blood 75:662-670(1990).
RN [23]
RN VARIANT HEMA CYS-391.
RX MEDLINE=90329422; PubMed=1973901;
RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
RT "CRM+ haemophilia A due to a missense mutation (372->Cys) at the
RT internal heavy chain thrombin cleavage site";
RL Br. J. Haematol. 75:73-77(1990).
RN [24]
RN VARIANTS HEMA PHE-1699 AND CYS-1708.
RX MEDLINE=90152691; PubMed=2105906;
RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
RA Kazazian H.H., Antonarakis S.E.;
RT "Characterization of mutations in the factor VIII gene by direct
RT sequencing of amplified genomic DNA";
RL Genomics 6:65-71(1990).
RN [25]
RN VARIANTS HEMA CYS-1728 AND ASP-1941.
RX MEDLINE=90169988; PubMed=2106480;
RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
RA Kazazian H.H.;
RT "Use of denaturing gradient gel electrophoresis to detect point
RT mutations in the factor VIII gene";

```

Query Match          100.0%; Score 155; DB 1; Length 2351;
Best Local Similarity 96.7%; Pred. No. 3.8e-14;
Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQLXCHSHQHDGME 30
    |||||
Db 311 TFLTAQTLLMDLGGQLXCHSHQHDGME 340

RESULT 2
FAB_MOUSE
ID FAB_MOUSE STANDARD; PRT; 2319 AA.
AC Q06194;
AT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
GN F8 OR CF8 OR F8C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=93300511; PubMed=8314577;
RA Elder B., Lakich D., Gitschier J.;
RT "Sequence of the murine factor VIII cDNA.";
RL Genomics 16:374-379(1993).
CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
CC ACTIVATED FORM, FACTOR XA.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05573; AAA37385.1;
CC PIR; A47004; A47004.
CC HSSP; P00451; 1CFG.
CC MGD; MGI:88383; F8.
CC InterPro; IPR001117; Cu-oxidase.
CC InterPro; IPR000421; FA58.C.
CC Pfam; PF00394; Cu-oxidase; 3.
CC Pfam; PF00754; F5_P8_type_C; 2.
CC SMART; SM00231; FA58C; 2.
CC PROSITE; PS01285; FA58C.1; 2.
CC PROSITE; PS01286; FA58C.2; 2.
CC PROSITE; PS00022; FA58C.3; 2.
CC PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
KW Signal; Glycoprotein; Sulfation.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2319 COAGULATION FACTOR VIII.
FT DOMAIN 20 349 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 207 349 PLASTOCYANIN-LIKE 1.
FT DOMAIN 399 730 F5/8 TYPE A 2.
FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
FT DOMAIN 760 1640 B.
FT DOMAIN 1683 2008 F5/8 TYPE A 3.
FT DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.

```

```

FT DOMAIN 1855 2008 PLASTOCYANIN-LIKE 6.
FT DOMAIN 2008 2156 F5/8 TYPE C 1.
FT DOMAIN 2161 392 F5/8 TYPE C 2.
FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT MOD_RES 367 367 SULFATION (BY SIMILARITY).
FT MOD_RES 737 737 SULFATION (BY SIMILARITY).
FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
FT MOD_RES 1669 1669 SULFATION (REQUIRED FOR VWF BINDING)
    (BY SIMILARITY).
FT MOD_RES 1687 1687 SULFATION (BY SIMILARITY).
FT DISULFID 173 199 PROBABLE.
FT DISULFID 547 573 PROBABLE.
FT DISULFID 1819 1845 PROBABLE.
FT DISULFID 2008 2156 BY SIMILARITY.
FT DISULFID 2161 2313 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 958 958 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1136 1136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1161 1161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1192 1192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 2319 AA; 266148 MW; 26054DE051DB2A01 CRC64;

```

```

Query Match          94.8%; Score 147; DB 1; Length 2319;
Best Local Similarity 90.0%; Pred. No. 5.3e-13;
Matches 27; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLGGQLXCHSHQHDGME 30
    |||||
Db 312 TFLTAQTLLMDLGGQLXCHSHQHDGME 341

RESULT 3
FAB_PIG
ID FAB_PIG STANDARD; PRT; 2133 AA.
AC P12263; O95243;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
GN F8 OR CF8.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Healey J.F., Lubin I.M., Lollar P.;

```

Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

[2]
SEQUENCE OF 705-1573 FROM N.A.
MEDLINE=86287369; PubMed=3016730;
Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C., Kaufman R.J.,
"A large region (approximately equal to 95 kDa) of human factor VIII is dispensable for in vitro procoagulant activity.";
Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
[3]
SEQUENCE OF 392-759 FROM N.A.
MEDLINE=94179260; PubMed=7510693;
Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
"Elimination of a major inhibitor epitope in factor VIII.";
J. Biol. Chem. 269:8639-8641(1994).
-!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE ACTIVATED FORM, FACTOR XA.
-!- SUBCELLULAR LOCATION: Extracellular.
-!- SIMILARITY: Contains 3 F5/8 type A domains.
-!- SIMILARITY: Contains 2 F5/8 type C domains.
-!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U49517; AB06705.1; -;
PIR; A25945; A25945.
PIR; T42763; T42763.
HSSP; P00451; 1CFG.
InterPro: IPR001117; Cu-oxidase.
InterPro: IPR000421; FA58_C.
Pfam; PF00394; Cu-oxidase; 3.
Pfam; PF00734; F5_F8_type_C; 2.
SMART; SM00231; FA58C; 2.
PROSITE; PS01285; FA58C_1; 2.
PROSITE; PS01286; FA58C_2; 2.
PROSITE; PS00022; FA58C_3; 2.
PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
Blood coagulation; Repeat; Plasma; Acute phase; Calcium; Signal; Glycoprotein; Sulfation.
SIGNAL 1 19 POTENTIAL.
CHAIN 20 2133 COAGULATION FACTOR VIII.
DOMAIN 20 357 F5/8 TYPE A 1.
DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
DOMAIN 207 357 PLASTOCYANIN-LIKE 2.
DOMAIN 399 730 F5/8 TYPE A 2.
DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
DOMAIN 760 1599 B.
DOMAIN 1495 1822 F5/8 TYPE A 3.
DOMAIN 1495 1659 PLASTOCYANIN-LIKE 5.
DOMAIN 1669 1822 PLASTOCYANIN-LIKE 6.
DOMAIN 1822 1970 F5/8 TYPE C 1.
DOMAIN 1975 2127 F5/8 TYPE C 2.
SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
SITE 1449 1450 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
MOD_RES 737 737 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
MOD_RES 738 738 SULFATION (BY SIMILARITY).
MOD_RES 742 742 SULFATION (BY SIMILARITY).
DISULFID 173 199 SULFATION (BY SIMILARITY).
DISULFID 547 573 PROBABLE.
DISULFID 1633 1659 PROBABLE.
DISULFID 1822 1970 BY SIMILARITY.
DISULFID 1975 2127 BY SIMILARITY.
CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 929 929 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 985 985 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1111 1111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1181 1181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1208 1208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1245 1245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1335 1335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1611 1611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1919 1919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 713 713 N -> M (IN REF. 2).
FT CONFLICT 734 734 I -> T (IN REF. 2).
FT CONFLICT 792 792 E -> Q (IN REF. 2).
FT CONFLICT 1133 1133 E -> F (IN REF. 2).
FT CONFLICT 1191 1191 I -> L (IN REF. 2).
FT CONFLICT 1209 1209 R -> F (IN REF. 2).
FT CONFLICT 1437 1437 C -> G (IN REF. 2).
FT CONFLICT 1456 1456 F -> R (IN REF. 2).
FT CONFLICT 1539 1539 F -> R (IN REF. 2).
FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
SQ SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;
Query Match 89.7%; Score 139; DB 1; Length 2133;
Best Local Similarity 86.7%; Pred. No. 6.9e-12;
Matches 26; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 TELTAQFLMLDGLGQFLXCHISSHQHDGME 30
Dy 312 TELTAQFLMLDGLGQFLXCHISSHHGME 341

RESULT 4
CERU_HUMAN
ID CERU_HUMAN STANDARD; PRT; 1065 AA.
AC P00450; Q14063;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ceruloplasmin precursor (EC 1.16.3.1) (ferroxidase).
GN CP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86259737; PubMed=2873574;
RA Koschinsky M.L., Funk W.D., van Oost B.A., McGillivray R.T.A.;
RT "Complete cDNA sequence of human ceruloplasmin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5086-5090(1986).
RN [2]
RP SEQUENCE OF 1-1006 FROM N.A.
RX MEDLINE=95217183; PubMed=7702601;
RA Daimon M., Yamatani K., Igarashi M., Fukase N., Kawanami T.,
Kato T., Tomimaga M., Sasaki H.;
RT "Fine structure of human ceruloplasmin gene.";
RL Biochem. Biophys. Res. Commun. 208:1028-1035(1995).
RN [3]
RP SEQUENCE OF 1-40; 549-599; 784-829 AND 919-952 FROM N.A.
RX MEDLINE=86275241; PubMed=3755405;
RA Mercer J.F.B., Grimes A.;
RT "Isolation of a human ceruloplasmin cDNA clone that includes the N-terminal leader sequence.";
RL FEBS Lett. 203:185-190(1986).
RN [4]
RP SEQUENCE OF 218-1065 FROM N.A.
RX MEDLINE=86205876; PubMed=3486416;
RA Yang F., Naylor S.L., Lum J.B., Cutshaw S., McCombs J.L.,
Naberhaus K.H., McGill J.R., Adrian G.S., Moore C.M., Barnett D.R.,


```

CC CC -!- SUBCELLULAR LOCATION: Nuclear.
CC CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M88116; AAA35401.1; -
CC CC DR PIR; JC1179; JC1179.
CC CC DR HSSP; P01106; 1A93.
CC CC DR TRANSTAC; T03433; -.
CC CC DR InterPro; IPR001092; HLH_basic.
CC CC DR InterPro; IPR003327; Myc-LZ.
CC CC DR InterPro; IPR002418; TF_Myc.
CC CC DR Pfam; PF00010; HLH; 1.
CC CC DR Pfam; PF02344; Myc-LZ; 1.
CC CC DR Pfam; PF01056; Myc_N_term; 1.
CC CC DR PRINTS; PR00044; LEUZIPPRMYC.
CC CC DR SMART; SM00353; HLH; 1.
CC CC DR PROSITE; PS00038; HLH.1; 1.
CC CC DR PROSITE; PS00888; HLH.2; 1.
CC CC KW Proto-oncogene; Nuclear protein; DNA-binding; Phosphorylation;
CC CC KW Transcription regulation; Activator; Glycoprotein.
CC CC FT DOMAIN 33 37
CC CC FT POLY-GLY.
CC CC FT DOMAIN 88 91
CC CC FT BASIC DOMAIN.
CC CC FT DOMAIN 353 366
CC CC FT HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
CC CC FT DOMAIN 367 406
CC CC FT LEUCINE-ZIPPER (POTENTIAL).
CC CC FT DOMAIN 412 433
CC CC FT LEUCINE-ZIPPER (POTENTIAL).
CC CC FT MOD_RES 58 58
CC CC FT PHOSPHORYLATION (BY SIMILARITY).
CC CC FT MOD_RES 62 62
CC CC FT PHOSPHORYLATION (BY SIMILARITY).
CC CC FT CARBOHYD 58 58
CC CC FT O-LINKED (GLCNAC) (BY SIMILARITY).
CC CC SQ SEQUENCE 438 AA; 48747 MW; BBE7E5141739DCE6 CRC64;

Query Match 32.3%; Score 50; DB 1; Length 438;
Best Local Similarity 58.3%; Pred. No. 9.4;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 16 LLXCHISSHQHD 27
Db 296 LKRCHVSTHOHN 307
: : : : :
: : : : :

RESULT 15
MYC_CANFA STANDARD; PRT; 439 AA.
ID MYC_CANFA
AC Q28350;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myc proto-oncogene protein (c-myc).
OS MYC.
OS Canis familiaris (Dog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OS Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=96257835; PubMed=8647923;
RA Pirson I., Coulonval K., Lamy F., Dumont J.E.;
RT "c-myc expression is controlled by the mitogenic cAMP-cascade in
RT thymocytes.;"
RL J. Cell. Physiol. 168:59-70(1996).
CC -!- FUNCTION: PARTICIPATES IN THE REGULATION OF GENE TRANSCRIPTION.
CC CC BINDS DNA BOTH IN A NON-SPECIFIC MANNER AND ALSO SPECIFICALLY TO
CC CC RECOGNIZES THE CORE SEQUENCE CAC[GATG]. SEEMS TO ACTIVATE THE
CC CC TRANSCRIPTION OF GROWTH-RELATED GENES.
CC CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER

```

```

CC BHLH PROTEIN. BINDS DNA AS AN HETERODIMER WITH MAX.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- DISEASE: OVEREXPRESSION OF C-MYC IS IMPLICATED IN THE ETIOLOGY OF
CC A VARIETY OF HEMATOPOIETIC TUMORS.
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X95367; CAA64654.1; -
CC HSSP: P01106; 1A93.
CC TRANSFAC: T03432; -
CC InterPro: IPR001092; HLH_Basic.
CC InterPro: IPR003327; MYC-LZ.
CC InterPro: IPR002416; TF_MYC.
CC Pfam: PF00010; HLH; 1.
CC Pfam: PF02344; MYC-LZ; 1.
CC Pfam: PF01056; MYC_N-term; 1.
CC PRINTS: PR00044; LEUZIPPRMYC.
CC SMART: SM00353; HLE; 1.
CC PROSITE: PS00038; HLH_1; 1.
CC PROSITE: PS00888; HLH_2; 1.
CC KW Proto-oncogene; Nuclear protein; DNA-binding; Phosphorylation;
CC Transcription regulation; Activator; Glycoprotein.
CC FT DOMAIN 33 37
CC FT DOMAIN 88 91
CC FT DNA_BIND 354 367
CC FT DOMAIN 368 407
CC FT DOMAIN 413 434
CC FT MOD_RES 58 58
CC FT MOD_RES 62 62
CC FT CARBOHYD 58 58
CC SQ SEQUENCE 439 AA; 48409 MW; 9A1D9F31C66C7C8B CRC64;

Query Match 32.3%; Score 50; DB 1; Length 439;
Best Local Similarity 58.3%; Pred. No. 9.4;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 16 LLXCHISSHQHD 27
Db 297 LARCHVSTHQHN 308

```

Search completed: October 21, 2003, 18:54:14
Job time : 11.25 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	τ				
1	147	94.8	355	11	O8BQ43		O8bq43 mus musculus
2	146	94.2	2343	6	O62730		O62730 canis famill
3	146	94.2	2343	6	O18806		O18806 canis famill
4	72	46.5	782	4	O75659		O75659 homo sapien
5	72	46.5	1104	4	O75180		O75180 homo sapien
6	72	46.5	1158	4	O9BQ57		O9bqs7 homo sapien
7	72	46.5	1158	4	O9C058		O9cqs7 homo sapien
8	68	43.9	1157	11	O920H8		O920h8 rattus norv
9	68	43.9	1157	11	O920Z4		O920z4 mus musculu
10	66	42.6	594	3	O96vT5		O96vt5 emericeila
11	66	42.6	847	11	O8C4S2		O8c4s2 mus musculu
12	64.5	41.6	1048	6	O9xtZ7		O9xtz7 ovis aries
13	63	40.6	503	11	O8BV37		O8bv37 mus musculu
14	62	40.0	1084	11	O9Jl97		O9jl97 rattus norv
15	61	39.4	669	2	O8KKW5		O8kkw5 rhizobium e
16	61	39.4	717	16	O8XPV6		O8xpv6 ralstonia s

SQ SEQUENCE 1104 AA; 124257 MW; 617C753F766AAL52 CRC64;
 Query Match 46.5%; Score 72; DB 4; Length 1104;
 Best Local Similarity 40.0%; Pred. No. 0.024;
 Matches 12; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLQFLXCHISSHQHDGME 30
 ||||: : : : : ||||: ||||
 DB 275 TFTAEMVPEPGTWLISCQVNSHFRDGMQ 304

RESULT 6
 Q9C057 PRELIMINARY; PRT; 1158 AA.
 ID Q9C057
 AC Q9C057;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hephaestin.
 GN HEPH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Syed B.A., Beaumont N., Evans R.W., Srai S.K.;
 RT "In silico cloning and comparative modelling of human Hephaestin."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ296162; CAC35365.2; -;
 DR HSP; P00450; 1KCW.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR002355; MultiCu_oxidase2.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 SQ SEQUENCE 1158 AA; 130450 MW; A3FDBF4D35F1E5AE CRC64;

Query Match 46.5%; Score 72; DB 4; Length 1158;
 Best Local Similarity 40.0%; Pred. No. 0.025;
 Matches 12; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLQFLXCHISSHQHDGME 30
 ||||: : : : : ||||: ||||
 DB 329 TFTAEMVPEPGTWLISCQVNSHFRDGMQ 358

RESULT 7
 Q9C058 PRELIMINARY; PRT; 1158 AA.
 ID Q9C058
 AC Q9C058;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2002 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hephaestin.
 GN HEPH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhao K.W.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF148860; AAK08131.1; -;
 DR EMBL; BC011561; AAH11561.1; -;
 DR HSP; P00450; 1KCW.
 DR Genew; HGNC:4866; HEPH.

DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR002355; MultiCu_oxidase2.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 SQ SEQUENCE 1158 AA; 130448 MW; CD032199E2E2868D CRC64;

Query Match 46.5%; Score 72; DB 4; Length 1158;
 Best Local Similarity 40.0%; Pred. No. 0.025;
 Matches 12; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLQFLXCHISSHQHDGME 30
 ||||: : : : : ||||: ||||
 DB 329 TFTAEMVPEPGTWLISCQVNSHFRDGMQ 358

RESULT 8
 Q920H8 PRELIMINARY; PRT; 1157 AA.
 ID Q920H8
 AC Q920H8;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hephaestin.
 GN HEPH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Frazer D.M., Vulpe C.D., McKie A.T., Wilkins S.J., Trinder D.,
 RA Clegborn G.J., Anderson G.J.;
 RT "Cloning and gastrointestinal expression of rat hephaestin:
 RT relationship to other iron transport proteins.";
 RL Am. J. Physiol. 281:G931-G939(2001).
 DR EMBL; AF246120; AAL08217.1; -;
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR002355; MultiCu_oxidase2.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 SQ SEQUENCE 1157 AA; 129593 MW; 0C626FA3E2F51DE2 CRC64;

Query Match 43.9%; Score 68; DB 11; Length 1157;
 Best Local Similarity 36.7%; Pred. No. 0.1;
 Matches 11; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLQFLXCHISSHQHDGME 30
 || : : : ||||: ||||
 DB 1027 TEEVEMVASNPGAWLMHCHVTDHVHAGME 1056

RESULT 9
 Q9Z0Z4 PRELIMINARY; PRT; 1157 AA.
 ID Q9Z0Z4
 AC Q9Z0Z4;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hephaestin.
 GN HEPH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=99140771; PubMed=9988272;
 RA Vulpe C.D., Kuo Y.M., Murphy T.L., Cowley L., Askwith C., Libina N.,
 RA Gitschier J., Anderson G.J.;

"Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse.";
 Nat. Genet. 21:195-199(1999).
 EMBL: AF082567; AAD16035.1; -.
 DR HSSP: P00450; 1KCW.

DR MGD: MGI:1332240; Heph.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; MultiCu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 SQ SEQUENCE 1157 AA; 129682 MW; 411BBD2A2BPEC2FD CRC64;

Query Match 43.9%; Score 68; DB 11; Length 1157;
 Best Local Similarity 36.7%; Pred. No. 0.1;
 Matches 11; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQFLXCHISSHQHGM 30
 || : : : : : |||
 Db 1027 TFEVEMVASNPGTWMHCHVTDHVRHAGME 1056

RESULT 10

Q96VT5 PRELIMINARY; PRT; 594 AA.
 ID Q96VT5
 AC Q96VT5;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Laccase precursor (EC 1.10.3.2).
 GN TILA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Scherer M.;
 RL Thesis (2001), Department of Microbiology,
 RL Philipps-Universitaet-Marburg, Marburg, Germany.
 DR EMBL: AJ305224; CAC59820.1; -.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; MultiCu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Oxidoreductase; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 594 LACCASE.
 SQ SEQUENCE 594 AA; 65816 MW; 1GAIE82F24F7F01C CRC64;

Query Match 42.6%; Score 66; DB 3; Length 594;
 Best Local Similarity 44.8%; Pred. No. 0.11;
 Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQFLXCHISSHQHGM 29
 || : : : : : |||
 Db 545 TWLALRYHVAVNPGPFLLECHIQHOSGGM 573

RESULT 11

Q8C4S2 PRELIMINARY; PRT; 847 AA.
 ID Q8C4S2
 AC Q8C4S2;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE HEPHAESTIN homologue.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL: AK081330; BAC38197.1; -.
 SQ SEQUENCE 847 AA; 95136 MW; 778CEC0617243402 CRC64;

Query Match 42.6%; Score 66; DB 11; Length 847;
 Best Local Similarity 36.7%; Pred. No. 0.16;
 Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 TFLTAQTLLMDLQFLXCHISSHQHGM 30
 ||| : : : : : |||
 Db 329 TFWTAEMVPQSGTWLISCEVNSHLRSGM 358

RESULT 12

Q9XT27 PRELIMINARY; PRT; 1048 AA.
 ID Q9XT27
 AC Q9XT27;
 DT 01-NOV-1999 (TREMELrel. 12, Created)
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE Ceruloplasmin.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Liver;
 RX MEDLINE=99384006; PubMed=10452945;
 RA Lockhart P.J., Mercer J.F.B.;
 RT "Cloning and expression analysis of the sheep ceruloplasmin cDNA.";
 RL Gene 236:251-257(1999).
 DR EMBL: AF134814; AAD41477.1; -.
 DR HSP: P00450; 1KCW.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; MultiCu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 SQ SEQUENCE 1048 AA; 119125 MW; 925F16D7B0549CBB CRC64;

Query Match 41.6%; Score 64.5; DB 6; Length 1048;
 Best Local Similarity 50.0%; Pred. No. 0.33;
 Matches 15; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY 4 TAOTLLM--DLGQFLXCHISSHQHGM 30
 ||||| : : : : : |||
 Db 1005 TYQILEMPKPTFGIWLHCHVTDHVRHAGME 1034

RESULT 13

Q8BV37 PRELIMINARY; PRT; 503 AA.
 ID Q8BV37
 AC Q8BV37;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Similar to GPI-anchored ceruloplasmin (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;

```

RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK080701; BAC37984.1; -.
FT NON_TER
SQ SEQUENCE 503 AA; 58067 MW; 042331F60B1853EB CRC64;

  Query Match      40.6%; Score 63; DB 11; Length 503;
  Best Local Similarity 37.9%; Pred. No. 0.27;
  Matches 11; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLGFLXCHISSHQHDGM 29
Db 459 TYQTELEFPQTPGTWLLCHVTHVHAGM 487

RESULT 14
Q9JL97 PRELIMINARY; PRT; 1084 AA.
AC Q9JL97
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GPI-anchored ceruloplasmin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague Dawley; TISSUE=Brain;
RX MEDLINE=20127919; PubMed=10660599;
RA Patel B.N., Dunn R.J., David S.;
RT "Alternative RNA Splicing Generates a Glycosylphosphatidylinositol-
RT anchored Form of Ceruloplasmin in Mammalian Brain.";
RL J. Biol. Chem. 275:4305-4310(2000).
DR EMBL; AF202115; AAF34175.1; -.
DR HSSP; P00450; 1RCW.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; MultiCu_oxidase2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
SQ SEQUENCE 1084 AA; 123749 MW; 95D65B4154257C55 CRC64;

  Query Match      40.0%; Score 62; DB 11; Length 1084;
  Best Local Similarity 37.9%; Pred. No. 0.84;
  Matches 11; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 TFLTAQTLMLDGLGFLXCHISSHQHDGM 29
Db 1016 TYQTELEFPQTPGTWLLCHVTHVHAGM 1044

RESULT 15
Q8KKW5 PRELIMINARY; PRT; 669 AA.
AC Q8KKW5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Yp076.
OS Rhizobium etli.
OG Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OC NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=CFN42;
RX MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
RT leguminosarum bv. phaseoli.";
RL J. Bacteriol. 173:2411-2419(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
RT of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase and RecA are required to exert
RT incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RL EMBL; U80928; AAM55015.1; -.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; MultiCu_oxidase2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 669 AA; 73046 MW; CE98B723C7B72996 CRC64;

  Query Match      39.4%; Score 61; DB 2; Length 669;
  Best Local Similarity 50.0%; Pred. No. 0.73;
  Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 12 LGQFLXCHISSHQHDGM 29
Db 632 IGEFVLHCHLHEDQGM 649

Search completed: October 21, 2003, 18:57:21
Job time : 47.75 secs

```


Misc_feature 711..746
 /note= "a spacer of the sequence
 SFSQNRHSTKQKFNATIPNDIEKTDPPWF
 AHTPMKIQNVSSDLMLL is inserted
 between domains A2 and A3"
 Domain 380..554
 /note= "plastocyanin-like domain 3"
 Domain 564..711
 /note= "plastocyanin-like domain 4"
 Domain 746..1073
 /note= "A3 domain"
 Domain 1073..1221
 /note= "C1 domain"
 Domain 1226..1378
 /note= "C2 domain"
 Cleavage-site 372..373
 /note= "by thrombin"
 Disulfide-bond 153..179
 /note= "probable"
 Disulfide-bond 528..554
 /note= "probable"
 Misc-difference 336
 /label= R336I
 /note= "wild type Arg replaced with Ile"
 Misc-difference 562
 /label= R562K
 /note= "wild type Arg replaced with Lys"
 Misc-difference 740
 /label= R740A
 /note= "wild type Arg replaced with Ala"

W09740145-AL.

30-OCT-1997.

24-APR-1997; 97WO-0506563.

15-MAY-1996; 96US-0017785.

24-APR-1996; 96US-0016117.

(UNMI) UNIV MICHIGAN.

Amato K, Kaufman RJ, Pipe SW;

WPI; 1997-535830/49.

Modified human pro-coagulant active factor VIII - can be
 administered to haemophiliacs, i.e. factor VIII replacement therapy
 Claim 18; Page -: 57pp; English.

The present sequence represents a novel pro-coagulant active factor
 VIII (FVIII) mutant protein, comprising a deletion of the B domain and
 von Willebrand factor binding site, mutations R336I, R562K and R740A and
 an addition of an amino acid sequence spacer between the A2 and A3
 domains. Factor VIII, along with calcium and phospholipid, acts as a
 cofactor for factor IXA, when it converts factor X to the activated form
 (factor XA). FVIII is the coagulation factor deficient in the
 X-chromosome-linked bleeding disorder haemophilia A. Several other
 mutant FVIII proteins have also been created (see AAW33222-29). The FVIII
 mutant F309S (AAW33225) is capable of recombinant secretion at higher
 levels than typically obtained with wild type FVIII and retains
 pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
 (AAW33223) are resistant to activated protein C (APC) cleavage. The FVIII
 mutant comprising a deletion of the B domain and von Willebrand factor
 binding site, a mutation at Arg740 and an addition of an amino acid
 sequence spacer between the A2 and A3 domains can form a more stable
 configuration, and have an approximate 5-fold increase in specific
 activity compared to purified wild type FVIII, while increasing their
 binding affinity to von Willebrand factor improves their stability.
 The FVIII proteins can be administered to haemophiliacs, i.e. FVIII
 replacement therapy, while the nucleic acid molecule can be used for
 gene therapy.

CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.
 XX
 SQ Sequence 1383 AA;
 Query Match 100.0%; Score 161; DB 18; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 7.6e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYKVDSCPEEPQLIMKNEEFAEDYDDDLT 30
 |||||
 Db 322 AYKVDSCPEEPQLIMKNEEFAEDYDDDLT 351
 |||||
 RESULT 2
 AAW33222
 ID AAW33222 standard; protein; 2332 AA.
 XX
 AC AAW33222;
 XX
 DT 30-APR-1998 (first entry)
 XX
 DE Procoagulant-active human factor VIII:C (FVIII) R336I mutant protein.
 XX
 KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PH Key
 FT Region Location/Qualifiers
 FT 1..1313
 FT /note= "factor VIIIA heavy chain"
 FT 1649..2332
 FT /note= "factor VIIIA light chain"
 FT 1..329
 FT /note= "A1 domain"
 FT 1..179
 FT /note= "plastocyanin-like domain 1"
 FT 187..329
 FT /note= "plastocyanin-like domain 2"
 FT 380..711
 FT /note= "A2 domain"
 FT 380..554
 FT /note= "plastocyanin-like domain 3"
 FT 564..711
 FT /note= "plastocyanin-like domain 4"
 FT 741..1648
 FT /note= "B domain"
 FT 1694..2021
 FT /note= "A3 domain"
 FT 1694..1858
 FT /note= "plastocyanin-like domain 5"
 FT 1868..2021
 FT /note= "plastocyanin-like domain 6"
 FT 2021..2169
 FT /note= "C1 domain"
 FT 2174..2326
 FT /note= "C2 domain"
 FT 372..373
 FT /note= "by thrombin"
 FT 740..741
 FT /note= "by thrombin"
 FT 1689..1690
 FT /note= "by thrombin"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Disulfide-bond 1832..1858

```

FT      /note= "probable"
FT      Misc-difference 336
FT      /label= R336I
FT      /note= "wild type Arg replaced with Ile"
PN      WO9740145-A1.
XX      30-OCT-1997.
XX      24-APR-1997; 97WO-US06563.
XX      15-MAY-1996; 96US-0017785.
PR      24-APR-1996; 96US-0016117.
XX      (UNMI ) UNIV MICHIGAN.
XX      Anano K, Kaufman RJ, Pipe SW;
XX      WPI; 1997-535830/49.
XX      Modified human pro-coagulant active factor VIII - can be
XX      administered to haemophiliacs, i.e. factor VIII replacement therapy
XX      Claim 10; Page -: 57pp; English.
XX      The present sequence represents a novel pro-coagulant active factor
XX      VIII (FVIII) mutant protein, comprising a mutation at Arg336, which is
XX      replaced with Ile. Factor VIII, along with calcium and phospholipid,
XX      acts as a cofactor for factor IXA, when it converts factor X to the
XX      activated form (factor IXA). FVIII is the coagulation factor deficient in
XX      the X-chromosome-linked bleeding disorder haemophilia A. Several other
XX      mutant FVIII proteins have also been created (see AAW33222-29). The
XX      FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
XX      higher levels than typically obtained with wild type FVIII and retains
XX      pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
XX      (AAW33223) are resistant to activated protein C (APC) cleavage. The
XX      FVIII mutant comprising a deletion of the B domain and von Willebrand
XX      factor binding site, a mutation at Arg740 and an addition of an amino
XX      acid sequence spacer between the A2 and A3 domains can form a more
XX      stable configuration, and have an approximate 5-fold increase in
XX      specific activity compared to purified wild type FVIII, while increasing
XX      their binding affinity to von Willebrand factor improves their
XX      stability. The FVIII proteins can be administered to haemophiliacs, i.e.
XX      FVIII replacement therapy, while the nucleic acid molecule can be used
XX      for gene therapy.
XX      note; this sequence does not appear in the specification; it was created
XX      using sequences from the given references.
XX      Sequence 2332 AA;
SQ      Query Match 100.0%; Score 161; DB 18; Length 2332;
        Best Local Similarity 100.0%; Pred No. 1.4e-13;
        Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AYKVDSCEEPQLMKNEEAEDYDDLT 30
        |||||
Db      322 AYKVDSCEEPQLMKNEEAEDYDDLT 351
        |||||
RESULT 3
AAW33224
ID      AAW33224 standard; protein; 2332 AA.
XX      AAW33224;
XX      30-APR-1998 (first entry)
XX      Procoagulant-active human factor VIII:C R336I/R562K mutant protein.
XX      Pro-coagulant active factor VIII; FVIII; haemophilia A;
XX      recombinant secretion; pro-coagulant activity; resistance;
XX      activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
XX      von Willebrand factor binding site; binding affinity;

```

```

KW      FVIII replacement therapy.
XX      Synthetic.
OS      Homo sapiens.
XX      Key
PH      Region
FT      Location/Qualifiers
FT      1..1313
FT      /note= "factor VIIIA heavy chain"
FT      1649..2332
FT      /note= "factor VIIIA light chain"
FT      1..329
FT      /note= "A1 domain"
FT      1..179
FT      /note= "plastocyanin-like domain 1"
FT      187..329
FT      /note= "plastocyanin-like domain 2"
FT      380..711
FT      /note= "A2 domain"
FT      380..554
FT      /note= "plastocyanin-like domain 3"
FT      564..711
FT      /note= "plastocyanin-like domain 4"
FT      741..1648
FT      /note= "B domain"
FT      1694..2021
FT      /note= "A3 domain"
FT      1694..1858
FT      /note= "plastocyanin-like domain 5"
FT      1868..2021
FT      /note= "plastocyanin-like domain 6"
FT      2021..2169
FT      /note= "C1 domain"
FT      2174..2332
FT      /note= "C2 domain"
FT      372..373
FT      /note= "by thrombin"
FT      740..741
FT      /note= "by thrombin"
FT      1689..1690
FT      /note= "by thrombin"
FT      Disulfide-bond 153..179
FT      /note= "probable"
FT      Disulfide-bond 528..554
FT      /note= "probable"
FT      Disulfide-bond 1832..1858
FT      /note= "probable"
FT      Misc-difference 336
FT      /label= R336I
FT      /note= "wild type Arg replaced with Ile"
FT      Misc-difference 562
FT      /label= R562K
FT      /note= "wild type Arg replaced with Lys"
XX      WO9740145-A1.
XX      30-OCT-1997.
XX      24-APR-1997; 97WO-US06563.
XX      15-MAY-1996; 96US-0017785.
PR      24-APR-1996; 96US-0016117.
XX      (UNMI ) UNIV MICHIGAN.
XX      Anano K, Kaufman RJ, Pipe SW;
XX      WPI; 1997-535830/49.
XX      Modified human pro-coagulant active factor VIII - can be
XX      administered to haemophiliacs, i.e. factor VIII replacement therapy
XX      Claim 10; Page -: 57pp; English.
XX

```

CC The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising the mutations Arg336Ile and
 CC Arg562Lys. Factor VIII, along with calcium and phospholipid,
 CC acts as a cofactor for factor IXA, when it converts factor X to the
 CC activated form (factor IXA). FVIII is the coagulation factor deficient in
 CC the X-chromosome-linked bleeding disorder haemophilia A. Several other
 CC mutant FVIII proteins have also been created (see AAW33222-29). The
 CC FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
 CC higher levels than typically obtained with wild type FVIII and retains
 CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
 CC (AAW33223) are resistant to activated protein C (APC) cleavage. The
 CC FVIII mutant comprising a deletion of the B domain and von Willebrand
 CC factor binding site, a mutation at Arg740 and an addition of an amino
 CC acid sequence spacer between the A2 and A3 domains can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

XX SQ Sequence 2332 AA;

Query Match 100.0%; Score 161; DB 18; Length 2332;
 Best Local Similarity 100.0%; Pred. NO. 1.4e-13;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQLMKNEAEEDYDDLLT 30
 |||||
 Db 322 AYVKVDSCEPQLMKNEAEEDYDDLLT 351

RESULT 4

AAW33226
 ID AAW33226 standard; protein; 2332 AA.

XX AC AAW33226;

DT 30-APR-1998 (first entry)

XX DE Procoagulant-active factor VIII:C F309S/R336I/R562K mutant protein.

KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.

XX OS Synthetic.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Region 1..1313
 FT /note= "factor VIIIA heavy chain"
 FT Region 1649..2332
 FT /note= "factor VIIIA light chain"
 FT Domain 1..329
 FT /note= "A1 domain"
 FT Domain 1..179
 FT /note= "plastocyanin-like domain 1"
 FT Domain 187..329
 FT /note= "plastocyanin-like domain 2"
 FT Domain 380..711
 FT /note= "A2 domain"
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Domain 741..1648
 FT /note= "B domain"
 FT Domain 1694..2021

FT Domain /note= "A3 domain"
 FT 1694..1858
 FT Domain /note= "plastocyanin-like domain 5"
 FT 1868..2021
 FT Domain /note= "plastocyanin-like domain 6"
 FT 2021..2169
 FT /note= "C1 domain"
 FT Domain 2174..2326
 FT /note= "C2 domain"
 FT 372..373
 FT Cleavage-site /note= "by thrombin"
 FT 740..741
 FT Cleavage-site /note= "by thrombin"
 FT 1689..1690
 FT Disulfide-bond /note= "by thrombin"
 FT 153..179
 FT Disulfide-bond /note= "probable"
 FT 528..554
 FT Disulfide-bond /note= "probable"
 FT 1822..1858
 FT Misc-difference /note= "probable"
 FT 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT 336
 FT Misc-difference /label= R336I
 FT /note= "wild type Arg replaced with Ile"
 FT 562
 FT Misc-difference /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 FT WC9740145-A1.
 XX PN 30-OCT-1997.
 XX PD
 XX XX
 XX PF 24-APR-1997; 97WO-US06563.
 XX PR 15-MAY-1996; 96US-0017785.
 XX PR 24-APR-1996; 96US-0016117.
 XX XX
 XX PA (UNMT) UNIV MICHIGAN.
 XX PI Amano K, Kaufman RJ, Pipe SW;
 XX DR WPI; 1997-535830/49.
 XX XX
 XX PT Modified human pro-coagulant active factor VIII - can be
 XX administered to haemophiliacs, i.e. factor VIII replacement therapy
 XX PS Claim 11; Page -: 57pp; English.
 XX XX
 CC The present sequence represents a novel pro-coagulant active factor
 CC VIII (FVIII) mutant protein, comprising the mutations Phe309Ser,
 CC Arg336Ile, and Arg562Lys. Factor VIII, with calcium and phospholipid,
 CC acts as a cofactor for factor IXA, when it converts factor X to the
 CC activated form (factor IXA). FVIII is the coagulation factor deficient in
 CC the X-chromosome-linked bleeding disorder haemophilia A. Several other
 CC mutant FVIII proteins have also been created (see AAW33222-29). The
 CC FVIII mutant F309S (AAW33225) is capable of recombinant secretion at
 CC higher levels than typically obtained with wild type FVIII and retains
 CC pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K
 CC (AAW33223) are resistant to activated protein C (APC) cleavage. The
 CC FVIII mutant comprising a deletion of the B domain and von Willebrand
 CC factor binding site, a mutation at Arg740 and an addition of an amino
 CC acid sequence spacer between the A2 and A3 domains can form a more
 CC stable configuration, and have an approximate 5-fold increase in
 CC specific activity compared to purified wild type FVIII, while increasing
 CC their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy.
 CC note: this sequence does not appear in the specification; it was created
 CC using sequences from the given references.

```

XX SQ Sequence 2332 AA;
Query Match 100.0%; Score 161; DB 18; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQPLMKNEEAEDYDDDLT 30
    |||||
Db 322 AYVKVDSCEPQPLMKNEEAEDYDDDLT 351

RESULT 5
AAR73019
ID AAR73019 standard; peptide; 372 AA.
XX
AC AAR73019;
XX
DT 25-MAR-2003 (updated)
DI 21-NOV-1995 (first entry)
XX
DE Human Factor-VIII N-terminal fragment.
XX
KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX
OS Homo sapiens.
XX
PN WO9513301-A1.
XX
PD 18-MAY-1995.
XX
PE 10-NOV-1994; 94WO-DK00424.
XX
PR 12-NOV-1993; 93DK-0001281.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Persson E;
XX
DR WPI; 1995-194038/25.
XX
PT Crosslinked Factor VIII polypeptide which is stable - is prep'd. using
PT bis(sulphosuccinimidy) suberate or disuccinimidy) suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting
PT activity
XX
PS Disclosure; Page 18; 36pp; English.
XX
CC This is the N-terminal fragment of human Factor-VIII which
CC may be crosslinked resulting in increased stability and retention
CC of high activity over extended periods of time after activation by
CC thrombin. The polypeptide is used to prevent or treat diseases
CC caused by the absence or deficiency of Factor-VIII in a subject
CC such as haemophilia.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 372 AA;
Query Match 95.7%; Score 154; DB 16; Length 372;
Best Local Similarity 96.7%; Pred. No. 1.5e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQPLMKNEEAEDYDDDLT 30
    |||||
Db 322 AYVKVDSCEPQPLMKNEEAEDYDDDLT 351

RESULT 6
AAR67709
ID AAR67709 standard; Protein; 400 AA.
XX
AC AAR67709;
XX

```

```

DT 25-MAR-2003 (updated)
DI 20-JUL-1995 (first entry)
XX
DE B-domain deleted Factor-VIII.
XX
KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;
KW adenovirus; vector.
XX
OS Homo sapiens.
XX
PN WO9429471-A1.
XX
PD 22-DEC-1994.
XX
PE 13-APR-1994; 94WO-US04075.
XX
PR 10-JUN-1993; 93US-0074920.
PR 25-MAR-1994; 94US-0218335.
XX
PA (GENE-) GENETIC THERAPY INC.
XX
PI Connelly S, Kaleko M, Smith T;
XX
DR WPI; 1995-036495/05.
DR N-PSDB; AAQ76016.
XX
PT New adenoviral vectors for treatment of haemophilia - contg. a
PT DNA sequence encoding a clotting factor, partic. Factor VIII or
PT Factor IX
XX
PS Disclosure; Fig. 17A-17C; 116pp; English.
XX
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
CC used to construct recombinant adenovirus vectors that produced
CC therapeutic levels of the clotting factor when administered to an
CC animal host, potentially providing hemophilia A gene therapy.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 400 AA;
Query Match 95.7%; Score 154; DB 16; Length 400;
Best Local Similarity 96.7%; Pred. No. 1.7e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQPLMKNEEAEDYDDDLT 30
    |||||
Db 340 AYVKVDSCEPQPLMKNEEAEDYDDDLT 369

RESULT 7
AAR74088
ID AAR74088 standard; protein; 720 AA.
XX
AC AAR74088;
XX
DT 25-MAR-2003 (updated)
DI 04-NOV-1995 (first entry)
XX
DE Factor-VIII heavy chain N-terminal fragment.
XX
KW human; Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX
OS Homo sapiens.
XX
PN WO9513300-A1.
XX
PD 18-MAY-1995.
XX
PR 10-NOV-1994; 94WO-DK00423.
XX
PR 12-NOV-1993; 93DK-0001280.
XX

```


PT comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX
 PS Disclosure; Page 11-14; 24pp; English.
 XX
 CC The new Factor-VIII/Factor-VIIIa derivative comprises a functional
 CC A2 domain in which amino acid residues, as indicated in the
 CC Features, are deleted or substituted by another amino acid so as to
 CC increase the overall positive charge. Asp-403, Asp-404, Asp-433,
 CC Asp-482 and Asp-500 are preferably substituted by Asn; Glu-434 and
 CC Glu-440 are preferably substituted by Gln. The new derivative has
 CC the same activity as the wild-type Factor-VIII but with improved
 CC stability (the activity is maintained for a longer period compared
 CC to the rapid decline of the activity of wt Factor-VIII). The new
 CC derivative can be used in a composition for treating diseases caused
 CC by an absence or deficiency of Factor-VIII, especially haemophilia.
 CC (Updated on 25-MAR-2003 to correct DR field.)
 XX
 SQ Sequence 740 AA;
 Query Match 95.7%; Score 154; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 3.5e-13;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 30
 DB 322 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 351
 RESULT 10
 AAR76961
 ID AAR76961 standard; protein; 740 AA.
 AC AAR76961;
 XX
 XX 09-MAR-1996 (first entry)
 DE Human Factor-VIII derivative.
 XX
 KW Factor-VIII; therapeutic; blood-clotting.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 692
 FT /label= absent or Ala, Thr, Ser, Gly or Asp
 FT Misc-difference 720
 FT /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729
 FT /label= absent or Val, Ala or Ile
 XX
 PN W09518827-Al.
 XX
 PD 13-JUL-1995.
 XX
 PF 06-JAN-1995; 95WO-DK00008.
 XX
 PR 07-JAN-1994; 94DK-0000032.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Ezban Rasmussen M, Nicolaisen EM, Persson E;
 XX
 PS WPI; 1995-255039/33.
 XX
 DR Novel factor VIII derivative used to treat haemophilia - and
 XX comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX
 PS Disclosure; Page 14-17; 30pp; English.
 XX
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is deleted or replaced with another amino acid

CC residue, preferably Ser (see AAR76962). Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX
 SQ Sequence 740 AA;
 Query Match 95.7%; Score 154; DB 16; Length 740;
 Best Local Similarity 96.7%; Pred. No. 3.5e-13;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 30
 DB 322 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 351
 RESULT 11
 AAR76962
 ID AAR76962 standard; protein; 740 AA.
 AC AAR76962;
 XX
 XX 09-MAR-1996 (first entry)
 DE Human Factor-VIII derivative.
 XX
 KW Factor-VIII; therapeutic; blood-clotting.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 692
 FT /label= Cys substituted by Ser
 FT Misc-difference 720
 FT /label= absent or Gln, Ser, Thr, Val or Ala
 FT Misc-difference 729
 FT /label= absent or Val, Ala or Ile
 XX
 PN W09518827-Al.
 XX
 PD 13-JUL-1995.
 XX
 PF 06-JAN-1995; 95WO-DK00008.
 XX
 PR 07-JAN-1994; 94DK-0000032.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Ezban Rasmussen M, Nicolaisen EM, Persson E;
 XX
 PS WPI; 1995-255039/33.
 XX
 DR Novel factor VIII derivative used to treat haemophilia - and
 XX comprises a functional A2 domain containing a mutation at one or
 PT more Cys residues.
 XX
 PS Disclosure; Page 18-20; 30pp; English.
 XX
 XX The new Factor-VIII derivative comprises a functional A2 domain in
 CC which Cys-692 is replaced with Ser. For other (less preferred)
 CC substitutions at this site, see AAR76961. Alternatively, Glu-720 and/or
 CC Tyr-729 are deleted or substituted with various amino acids (as in
 CC the Features). The new derivative has the same activity as the wild-
 CC type Factor-VIII but with improved stability (the activity is
 CC maintained for a longer period compared to the rapid decline of the
 CC activity of wt Factor-VIII). The new derivative can be used in a
 CC composition for treating diseases caused by an absence or deficiency
 CC of Factor-VIII, especially haemophilia.
 XX

```

SQ Sequence 740 AA;
  Query Match 95.7%; Score 154; DB 16; Length 740;
  Best Local Similarity 96.7%; Pred. No. 3.5e-13;
  Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 30
  |||||
Db 322 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 351

RESULT 12
AAR73021
ID AAR73021 standard; peptide; 740 AA.
XX
AC
XX
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
DE Human Factor-VIII N-terminal fragment.
XX
KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX
OS Homo sapiens.
XX
PN WO9513301-Al.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK00424.
XX
PR 12-NOV-1993; 93DK-0001281.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Persson E;
XX
DR WPI; 1995-194038/25.
XX
PT Crosslinked factor VIII polypeptide which is stable - is prep'd. using
PT bis(sulphosuccinimidylyl) suberate or disuccinimidylyl suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting
PT activity
XX
PS Disclosure; Page 21; 36pp; English.
XX
CC This is the N-terminal fragment of human Factor-VIII which may be
CC crosslinked resulting in increased stability and retention of high
CC activity over extended periods of time after activation by thrombin.
CC The polypeptide is used to prevent or treat diseases caused by the
CC absence or deficiency of Factor-VIII in a subject such as
CC haemophilia.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 740 AA;
  Query Match 95.7%; Score 154; DB 16; Length 740;
  Best Local Similarity 96.7%; Pred. No. 3.5e-13;
  Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 30
  |||||
Db 322 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 351

RESULT 13
AAR74090
ID AAR74090 standard; protein; 740 AA.
XX
AC AAR74090;
XX
DT 25-MAR-2003 (updated)

```

```

DT 04-NOV-1995 (first entry)
XX
XX Factor-VIII heavy chain N-terminal fragment.
XX
XX human; Factor VIII; heavy chain; N-terminal fragment;
KW thrombin cleavage; blood-clotting.
XX
XX Homo sapiens.
XX
PN WO9513300-Al.
XX
XX 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK00423.
XX
PR 12-NOV-1993; 93DK-0001280.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Ezban Rasmussen M, Kjalke M;
XX
DR WPI; 1995-194037/25.
XX
PT Factor VIII polypeptide(s) comprising a heavy chain shorter than native
PT Al-A2 domain - are easier to produce recombinantly and retain coagulant
PT activity, may be used to treat patients who have developed antibodies to
PT C-terminal epitope(s) of Factor VIII
XX
PS Disclosure; Page 30-32; 51pp; English.
XX
CC The sequence represents N-terminal residues 1-740 of a human Factor-
CC VIII heavy chain. The sequence contains entire A1 and A2 domains,
CC and truncated forms (1-720 in (AAR74088) and 1-729 in (AAR74090))
CC may be produced by treatment with a protease, e.g. thrombin. The
CC C-terminally truncated fragments have the same coagulant specific
CC activity as full-length Factor-VIII, and may be produced
CC recombinantly to reduce production costs and improve safety, giving
CC higher production levels and stability than for the full-length form.
CC The fragments may be used to treat patients who have developed
CC antibodies against epitopes in the C-terminal part of the heavy chain.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 740 AA;
  Query Match 95.7%; Score 154; DB 16; Length 740;
  Best Local Similarity 96.7%; Pred. No. 3.5e-13;
  Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 30
  |||||
Db 322 AYVKVDSCEPPEPQLMKNEEAEDYDDDLT 351

RESULT 14
AAR75982
ID AAR75982 standard; protein; 740 AA.
XX
AC AAR75982;
XX
DT 21-FEB-1996 (first entry)
XX
DE Human factor VIII A2-domain derivative.
XX
KW Factor VIII; human; haemophilia; thrombin; protein C; plasmin;
KW serine protease; recombination; therapy; deficiency.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 720 /label= Gln, Ser, Thr, Val, Ala
FT Misc-difference 729 /label= Val, Ala, Ile
FT

```



```

XX WO9518828-A1.
XX
XX 13-JUL-1995.
XX
XX 06-JAN-1995; 95WO-DK00009.
XX
XX 07-JAN-1994; 94DK-0000031.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Ezban Rasmussen M, Nicolaisen EM, Persson E;
XX WPI; 1995-255040/33.
XX
XX Novel factor VIII derivative with resistance to enzymatic cleavage -
XX PT and comprises a functional A2 domain where Glu720 and/or Tyr729 is
XX PT deleted/substituted.
XX
XX Claim 1; Page 11-14; 26pp; English.
XX
XX This sequence represents the A2 domain of a human factor VIII derivative.
XX CC Factor VIII is a large glycoprotein which is present in plasma at low
XX CC concentrations. Factor VIII is an essential part of the clotting
XX CC reaction in response to a wound. Factor VIII is susceptible to cleavage
XX CC by thrombin, activated protein C, plasmin, and other serine proteases.
XX CC Full length factor VIII consists of three repeats of the A-domain, a
XX CC B-domain and 2 repeats of the C-domain. Active factor VIII has the A1
XX CC domain cleaved off. Factor VIII is too unstable for use in recombinant
XX CC techniques. Factor VIII containing this sequence has improved stability
XX CC and shows resistance against enzymatic activity present in mammalian
XX CC cells. This means that factor VIII containing this sequence can be used
XX CC in recombinant techniques. A factor VIII derivative can be used for
XX CC treating diseases caused by an absence or deficiency of factor VIII (in
XX CC the same way as normal factor VIII) e.g. haemophilia. The advantage with
XX CC using a recombinant factor VIII also includes no need for lots of donors
XX CC in order to get a sufficient amount. Also, there is no long purification
XX CC process, and there is no risk of transmission of blood-borne diseases
XX CC such as HIV.
XX
XX SQ Sequence 740 AA;
XX
XX Query Match 95.7%; Score 154; DB 16; Length 740;
XX Best Local Similarity 96.7%; Pred. No. 3.5e-13;
XX Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYKVDSCPEEPOLIMKNEEADYDDDLT 30
DB 322 AYKVDSCPEEPOLIMKNEEADYDDDLT 351

RESULT 15
AAP71139
ID AAP71139 standard; Protein; 1014 AA.
XX
XX AAP71139;
XX
XX 25-MAR-2003 (updated)
XX DT 14-MAY-1991 (first entry)
XX
XX Facor VIII:c variant.
XX
XX Factor VIII:c; variant; proteolysis; resistance;
XX KW pro coagulation activity.
XX
XX Homo sapiens.
XX
XX WO8707144-A.
XX
XX 03-DEC-1987.
XX
XX 29-MAY-1987; 87WO-US01299.
XX

```

```

PR 29-MAY-1986; 86US-0868410.
PR 18-NOV-1986; 86US-0932767.
XX 09-DEC-1986; 86US-0939658.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Kaufman RJ, Pittman DD, Toole JJ;
XX
XX WPI; 1987-348539/49.
XX
XX New deletion and replacement variants of factor VIII:c - resistant
XX PT to proteolysis but retaining pro coagulant activity, and mew DNA
XX PT coding sequences.
XX
XX PS Disclosure; Page 1; 42pp; English.
XX
XX The full-length human factor VIII:c cDNA has been set forth in
XX CC detail in WO8501961. This sequences is an example of the formula
XX CC A-X-B, wherein A-Ala1-Arg372, B-Ser1690-Tyr2332 and X=0-1316
XX CC amino acids substantially, duplicative of sequences of amino acids
XX CC within the sequence Arg372-Ser1690 of the full-length sequence.
XX CC Here X=0; producing a Arg371-Ser1690 fusion protein.
XX CC One or more deletions or replacements at Arg 220, 226, 279, 282,
XX CC 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
XX CC variants which are more resistant to specific proteolytic cleavage
XX CC compared with natural factor VIII:c. Pro-coagulant activity and
XX CC thrombin activability have been retained.
XX CC See also AAP711726-29.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 1014 AA;
XX
XX Query Match 95.7%; Score 154; DB 8; Length 1014;
XX Best Local Similarity 96.7%; Pred. No. 5.2e-13;
XX Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYKVDSCPEEPOLIMKNEEADYDDDLT 30
DB 322 AYKVDSCPEEPOLIMKNEEADYDDDLT 351

Search completed: October 21, 2003, 18:53:30
Job time : 49.5 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:50:07 ; Search time 16 Seconds
(without alignments)
79.333 Million cell updates/sec

Title: FVIII_ARG355I
Perfect score: 161
Sequence: 1 AYKVDSCPEEPQLMKNEEAEDYDDDLT 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154	95.7	1438	4	US-09-209-916-1
2	154	95.7	1471	1	US-08-683-839B-3
3	154	95.7	1661	2	US-08-882-083-2
4	154	95.7	1661	2	US-08-558-107-2
5	154	95.7	1661	3	US-09-243-539-2
6	154	95.7	2332	1	US-07-864-004B-4
7	154	95.7	2332	1	US-08-251-937A-4
8	154	95.7	2332	1	US-08-212-133A-2
9	154	95.7	2332	1	US-08-276-594A-2
10	154	95.7	2332	1	US-08-474-503-2
11	154	95.7	2332	2	US-08-670-707A-2
12	154	95.7	2332	3	US-09-037-601-2
13	154	95.7	2332	3	US-09-324-867-3
14	154	95.7	2332	4	US-09-315-179-2
15	154	95.7	2332	4	US-09-523-656-2
16	154	95.7	2332	5	PCT-US93-03275-4
17	154	95.7	2332	5	PCT-US94-13200-2
18	154	95.7	2351	1	US-08-121-202-2
19	154	95.7	2351	1	US-08-366-851A-2
20	154	95.7	2351	6	5171844-2
21	154	95.7	2351	6	5422260-1
22	120.5	74.8	2343	3	US-09-324-867-2
23	113.5	70.5	2304	3	US-09-324-867-4
24	113.5	70.5	2319	1	US-08-212-133A-8
25	113.5	70.5	2319	1	US-08-474-503-6
26	113.5	70.5	2319	2	US-08-670-707A-6
27	113.5	70.5	2319	3	US-09-037-601-6

28	113.5	70.5	2319	4	US-09-315-179-6	Sequence 6, Appli
29	113.5	70.5	2319	4	US-09-523-656-28	Sequence 28, Appli
30	113.5	70.5	2319	5	PCT-US94-13200-6	Sequence 6, Appli
31	90.5	56.2	541	1	US-08-121-202-4	Sequence 4, Appli
32	90.5	56.2	1443	2	US-08-670-707A-39	Sequence 39, Appli
33	90.5	56.2	1443	3	US-09-037-601-39	Sequence 39, Appli
34	90.5	56.2	1443	4	US-09-315-179-39	Sequence 39, Appli
35	90.5	56.2	1467	4	US-09-523-656-38	Sequence 38, Appli
36	90.5	56.2	2115	3	US-09-324-867-5	Sequence 5, Appli
37	90.5	56.2	2133	3	US-08-670-707A-37	Sequence 37, Appli
38	90.5	56.2	2133	3	US-09-037-601-37	Sequence 37, Appli
39	90.5	56.2	2133	4	US-09-315-179-37	Sequence 30, Appli
40	90.5	56.2	2133	4	US-09-523-656-30	Sequence 15, Appli
41	84	52.2	117	3	US-08-722-240-15	Sequence 20, Appli
42	50	31.1	1170	1	US-08-313-288B-20	Sequence 5413, Ap
43	47.5	29.5	87	4	US-09-134-001C-5413	Sequence 5, Appli
44	47	29.2	554	1	US-08-440-377A-5	Sequence 5, Appli
45	47	29.2	554	2	US-08-687-852-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

Query Match 95.7%; Score 154; DB 4; Length 1438;
Best Local Similarity 96.7%; Pred. No. 6,9e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYKVDSCPEEPQLMKNEEAEDYDDDLT 30
|||||
DB 322 AYKVDSCPEEPQLMKNEEAEDYDDDLT 351

RESULT 2
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: ILL, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences to Increase Expression of
; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-839B-3

```

```

Query Match          95.7%; Score 154; DB 1; Length 1471;
Best Local Similarity 96.7%; Pred. No. 7.1e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 30
   ||||||||||| |||||||||||
Db 341 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 370

```

```

RESULT 3
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-882-083-2

```

```

Query Match          95.7%; Score 154; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 8.2e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 30
   ||||||||||| |||||||||||
Db 341 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 370

```

```

RESULT 4
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-107-2

```

```

Query Match          95.7%; Score 154; DB 2; Length 1661;
Best Local Similarity 96.7%; Pred. No. 8.2e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 30
   ||||||||||| |||||||||||
Db 341 AYVKVDSCEPPEQLIMKNEEAEDYDDDLT 370

```

```

RESULT 5
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

```
;; ADDRESSEE: Foley & Iardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/243,539
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/558,107
;; FILING DATE: 13-NOV-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ISACSON, John P.
;; REGISTRATION NUMBER: 33,715
;; REFERENCE/DOCKET NUMBER: 30472/212
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1661 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-243-539-2

Query Match 95.7%; Score 154; DB 3; Length 1661;
Best Local Similarity 96.7%; Pred. No. 8.2e-14;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSPEEPQLMKNEEAEDYDDDLT 30
DB 341 AYVKVDSPEEPQLMKNEEAEDYDDDLT 370

RESULT 6
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
```

```
;; TELEPHONE: 404-815-6508
;; TELEFAX: 404-815-6555
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2332 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapien
;; TISSUE TYPE: Liver cDNA sequence
US-07-864-004B-4

Query Match 95.7%; Score 154; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSPEEPQLMKNEEAEDYDDDLT 30
DB 322 AYVKVDSPEEPQLMKNEEAEDYDDDLT 351

RESULT 7
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
```

```

; TISSUE TYPE: Liver cdna sequence
US-08-251-937A-4

Query Match          95.7%; Score 154; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQLMKNEAEYDDDLT 30
    |||||
Db 322 AYVKVDSCEPQLMKNEAEYDDDLT 351

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Iollar, John S.
; APPLICANT: RUNGE, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
US-08-212-133A-2

Query Match          95.7%; Score 154; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQLMKNEAEYDDDLT 30
    |||||
Db 322 AYVKVDSCEPQLMKNEAEYDDDLT 351

RESULT 9
US-08-276-594A-2

; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-276-594A-2

Query Match          95.7%; Score 154; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPQLMKNEAEYDDDLT 30
    |||||
Db 322 AYVKVDSCEPQLMKNEAEYDDDLT 351

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:

```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
US-08-474-503-2
```

```
Query Match 95.7%; Score 154; DB 1; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 AYVKVDSQPEEPQLMKNEEAEDYDDDLT 30
|||||
DB 322 AYVKVDSQPEEPQLMKNEEAEDYDDDLT 351
```

```
RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-08-670-707A-2
```

```
Query Match 95.7%; Score 154; DB 2; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 AYVKVDSQPEEPQLMKNEEAEDYDDDLT 30
|||||
DB 322 AYVKVDSQPEEPQLMKNEEAEDYDDDLT 351
```

```
RESULT 12
US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
US-09-037-601-2

Query Match 95.7%; Score 154; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 30
Db 322 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 351

RESULT 13

US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
; APPLICANT: Lilliecrap, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632lev. Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669.001002/JAG/BJD
; CURRENT APPLICATION NUMBER: US/09/324,867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-059
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-324-867-3

Query Match 95.7%; Score 154; DB 3; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 30
Db 323 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 352

RESULT 14

US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200

; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-179-2

Query Match 95.7%; Score 154; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 30
Db 322 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 351

RESULT 15

US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

Query Match 95.7%; Score 154; DB 4; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-13;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 30
Db 322 AYVKVDSCEPPEPQLRMKNNEEAEDYDDDLT 351

Search completed: October 21, 2003, 18:59:48
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 18:53:43 ; Search time 21.25 Seconds
(without alignments)
236.415 Million cell updates/sec

Title: FVIII_ARG355I
Perfect score: 161
Sequence: 1 AYVKVDSCEPPEQLMKNEEADYDDLT 30

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154	95.7	1438	14	US-10-006-091-1
2	154	95.7	1438	14	US-10-047-257-1
3	154	95.7	1438	15	US-10-225-900-1
4	154	95.7	1471	14	US-10-095-718-2
5	154	95.7	2332	10	US-09-957-641-2
6	154	95.7	2332	12	US-10-131-510A-2
7	154	95.7	2332	15	US-10-187-319-2
8	154	95.7	2351	12	US-10-133-907-4
9	154	95.7	2351	15	US-10-132-829-4
10	154	95.7	2351	15	US-10-172-712-27
11	130	80.7	29	9	US-09-853-080-14
12	120.5	74.8	1431	14	US-10-095-718-4
13	113.5	70.5	2319	12	US-10-131-510A-6
14	113.5	70.5	2319	15	US-10-187-319-6
15	90.5	56.2	1443	12	US-10-131-510A-39

16 90.5 56.2 1443 15 US-10-187-319-39 Sequence 39, Appli
17 90.5 56.2 2133 12 US-10-131-510A-37 Sequence 37, Appli
18 90.5 56.2 2133 15 US-10-187-319-37 Sequence 37, Appli
19 81 50.3 16 7 US-08-765-837-13 Sequence 13, Appli
20 52 32.3 192 15 US-10-106-698-5078 Sequence 5078, Ap
21 52 32.3 649 10 US-09-983-204-17 Sequence 17, Appli
22 52 32.3 649 14 US-10-133-157-6 Sequence 6, Appli
23 51.5 32.0 246 12 US-10-032-585-7042 Sequence 7042, Ap
24 51.5 32.0 520 9 US-09-213-678-2 Sequence 2, Appli
25 51.5 32.0 520 12 US-10-032-585-7035 Sequence 7035, Ap
26 50 31.1 37 12 US-10-058-053A-287 Sequence 287, App
27 50 31.1 37 12 US-10-058-053A-288 Sequence 288, App
28 50 31.1 37 12 US-10-058-053A-289 Sequence 289, App
29 50 31.1 109 12 US-10-058-053A-143 Sequence 143, App
30 50 31.1 110 12 US-10-058-053A-167 Sequence 167, App
31 50 31.1 110 12 US-10-058-053A-173 Sequence 173, App
32 50 31.1 427 12 US-10-032-585-7386 Sequence 7386, Ap
33 50 31.1 1152 10 US-09-919-603-1 Sequence 1, Appli
34 50 31.1 1170 12 US-10-021-660-114 Sequence 114, App
35 50 31.1 1170 12 US-10-008-093-2 Sequence 2, Appli
36 50 31.1 1170 15 US-10-020-141-12 Sequence 12, Appli
37 50 31.1 1170 15 US-10-017-721-2 Sequence 2, Appli
38 48 29.8 164 12 US-10-032-585-7315 Sequence 7315, Ap
39 48 29.8 304 10 US-09-866-582-27 Sequence 27, Appli
40 48 29.8 480 9 US-09-734-569-172 Sequence 172, App
41 48 29.8 532 12 US-10-032-189-36 Sequence 36, Appli
42 48 29.8 553 9 US-09-738-878-2 Sequence 2, Appli
43 48 29.8 553 14 US-10-163-381-2 Sequence 2, Appli
44 48 29.8 669 15 US-10-083-357-1336 Sequence 1336, Ap
45 47.5 29.5 69 9 US-09-864-761-37751 Sequence 37751, A

ALIGNMENTS

RESULT 1
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match 95.7%; Score 154; DB 14; Length 1438;
Best Local Similarity 96.7%; Pred. No. 6.9e-12;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYVKVDSCEPPEQLMKNEEADYDDLT 30
|||||
Db 322 AYVKVDSCEPPEQLMKNEEADYDDLT 351

RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:


```

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-510A-2

```

```

Query Match          95.7%; Score 154; DB 12; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-11;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  AYYKVDSCPEEPQLMKNEEAEDYDDDLT 30
        ||||||| ||||| ||||| ||||| |||||
Db      322 AYYKVDSCPEEPQLMKNEEAEDYDDDLT 351

```

```

RESULT 7
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/523,656
; FILING DATE: 2000-03-10
; APPLICATION NUMBER: US 09/037,601
; FILING DATE: 1998-03-10
; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: <Unknown>
; MOLECULE TYPE: Protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

```

```

Query Match          95.7%; Score 154; DB 15; Length 2332;
Best Local Similarity 96.7%; Pred. No. 1.2e-11;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  AYYKVDSCPEEPQLMKNEEAEDYDDDLT 30
        ||||||| ||||| ||||| ||||| |||||
Db      322 AYYKVDSCPEEPQLMKNEEAEDYDDDLT 351

```

```

RESULT 8
US-10-133-907-4
; Sequence 4, Application US/10133907
; Publication No. US20030195223A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pat170
; CURRENT APPLICATION NUMBER: US/10/133,907
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-907-4

```

```

Query Match          95.7%; Score 154; DB 12; Length 2351;
Best Local Similarity 96.7%; Pred. No. 1.2e-11;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  AYYKVDSCPEEPQLMKNEEAEDYDDDLT 30
        ||||||| ||||| ||||| ||||| |||||
Db      341 AYYKVDSCPEEPQLMKNEEAEDYDDDLT 370

```

```

RESULT 9
US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pat170
; CURRENT APPLICATION NUMBER: US/10/132,829
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-4

```

Query Match 95.7%; Score 154; DB 15; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 1.2e-11;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYKVDSCPEPQLMKNEEAEDYDDDLT 30
 |||||
 Db 341 AYKVDSCPEPQLMKNEEAEDYDDDLT 370

RESULT 10

US-10-172-712-27
 ; Sequence 27, Application US/10172712
 ; Publication No. US20030125232A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRIFFIN, JOHN H.
 ; APPLICANT: GALE, ANDREW J.
 ; APPLICANT: GETZOFF, ELIZABETH D.
 ; APPLICANT: PELLEQUER, JEAN-LUC
 ; TITLE OF INVENTION: STABILIZED PROTEINS WITH ENGINEERED DISULFIDE BONDS
 ; FILE REFERENCE: 4198-4001US1
 ; CURRENT APPLICATION NUMBER: US/10/172,712
 ; CURRENT FILING DATE: 2002-09-30
 ; PRIOR APPLICATION NUMBER: 60/238,578
 ; PRIOR FILING DATE: 2001-06-14
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 27
 ; LENGTH: 2351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-172-712-27

Query Match 95.7%; Score 154; DB 15; Length 2351;
 Best Local Similarity 96.7%; Pred. No. 1.2e-11;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYKVDSCPEPQLMKNEEAEDYDDDLT 30
 |||||
 Db 341 AYKVDSCPEPQLMKNEEAEDYDDDLT 370

RESULT 11

US-09-853-080-14
 ; Sequence 14, Application US/09853080
 ; Patent No. US20020068303A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Laub, Ruth
 ; APPLICANT: Di Giambattista, Mario
 ; TITLE OF INVENTION: ANTIGENIC POLYPEPTIDE SEQUENCES OF FACTOR
 ; FILE REFERENCE: VANMA48.001CPI
 ; CURRENT APPLICATION NUMBER: US/09/853,080
 ; CURRENT FILING DATE: 2001-05-09
 ; PRIOR APPLICATION NUMBER: US 08/765,837
 ; PRIOR FILING DATE: 1999-09-07
 ; PRIOR APPLICATION NUMBER: PCT/BE95/00068
 ; PRIOR FILING DATE: 1995-07-14
 ; PRIOR APPLICATION NUMBER: BE 9400666
 ; PRIOR FILING DATE: 1994-07-14
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 29
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: epitope Asp 327 to Met 355 of A1 domain of Factor
 US-09-853-080-14

Query Match 80.7%; Score 130; DB 9; Length 29;
 Best Local Similarity 96.0%; Pred. No. 1.3e-10;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 DSCPEPQLMKNEEAEDYDDDLT 30
 |||||
 Db 1 DSCPEPQLMKNEEAEDYDDDLT 25

RESULT 12

US-10-095-718-4
 ; Sequence 4, Application US/10095718
 ; Publication No. US20020131956A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walsh, Christopher
 ; APPLICANT: Chao, Hengjun
 ; APPLICANT: Burstein, Haim
 ; APPLICANT: Lynch, Carmel
 ; APPLICANT: Stepan, Tony
 ; APPLICANT: Munson, Keith
 ; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
 ; FILE REFERENCE: 35052/204375
 ; CURRENT APPLICATION NUMBER: US/10/095,718
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/689,430
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/158,780
 ; PRIOR FILING DATE: 1999-10-12
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 1431
 ; TYPE: PRT
 ; ORGANISM: canine B-domain deleted factor VIII
 US-10-095-718-4

Query Match 74.8%; Score 120.5; DB 14; Length 1431;
 Best Local Similarity 82.8%; Pred. No. 2e-07;
 Matches 24; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 AYKVDSCPEPQLMKNEEAEDYDDDL 29
 |||||
 Db 336 AYKVDSCPEPQLMKNEED-KDYDDGL 363

RESULT 13

US-10-131-510A-6
 ; Sequence 6, Application US/10131510A
 ; Publication No. US20030166536A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S
 ; TITLE OF INVENTION: Modified Factor VIII
 ; FILE REFERENCE: 75-95J
 ; CURRENT APPLICATION NUMBER: US/10/131,510A
 ; CURRENT FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: U.S. 09/315,179
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: U.S. 09/037,601
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: U.S. 08/670,707
 ; PRIOR FILING DATE: 1996-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US97/11155
 ; PRIOR FILING DATE: 1997-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/13200
 ; PRIOR FILING DATE: 1994-11-15
 ; PRIOR APPLICATION NUMBER: U.S. 08/212,133
 ; PRIOR FILING DATE: 1994-03-11
 ; PRIOR APPLICATION NUMBER: U.S. 07/864,004
 ; PRIOR FILING DATE: 1992-04-07
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 2319
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-10-131-510A-6

Query Match 70.5%; Score 113.5; DB 12; Length 2319;
Best Local Similarity 80.0%; Pred. No. 2.9e-06;
Matches 24; Conservative 0; Mismatches 5; Indels 1;

Qy 1 AYKVDSCPEEPQLIMK-NNEEAEDYDDDL 29
 ++++++++ + ++++++++
 Db 342 AYKVDSCPEESOWOKNNNEEMEDYDDDL 371

RESULT 14

```

US-10-187-319-6
; GENERAL INFORMATION:
;
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/187,319
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>

```

```

;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-187-319-6

Query Match          70.5%;   score 113.5;   DB 15;   Length 2319;
Best Local Similarity 80.0%;   Pred. No. 2.9e-05;
Matches 24; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

QY 1 AYKVDSCPEEPQLIMK -NNEAEDYDDDL 29
 ||||| 1 1 |||||
 Db 342 AYKVDSCPEESOWKKNNEEMEDYDDDL 371

RESULT 15

US-10-131-510A-39
; Sequence 39, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S

```

/ TITLE OF INVENTION: Modified Factor VIII
/
/ FILE REFERENCE: 75-95J
/
/ CURRENT APPLICATION NUMBER: US/10/131,510A
/
/ CURRENT FILING DATE: 2002-07-10
/
/ PRIOR APPLICATION NUMBER: U.S. 09/315,179
/
/ PRIOR FILING DATE: 1999-05-20
/
/ PRIOR APPLICATION NUMBER: U.S. 09/037,601
/
/ PRIOR FILING DATE: 1998-03-10
/
/ PRIOR APPLICATION NUMBER: U.S. 08/670,707
/
/ PRIOR FILING DATE: 1996-06-26
/
/ PRIOR APPLICATION NUMBER: PCT/US97/11155
/
/ PRIOR FILING DATE: 1997-06-26
/
/ PRIOR APPLICATION NUMBER: PCT/US94/13200
/
/ PRIOR FILING DATE: 1994-11-15
/
/ PRIOR APPLICATION NUMBER: U.S. 08/212,133
/
/ PRIOR FILING DATE: 1994-03-11
/
/ PRIOR APPLICATION NUMBER: U.S. 07/864,004
/
/ PRIOR FILING DATE: 1992-04-07
/
/ NUMBER OF SEQ ID NOS: 40
/
/ SOFTWARE: PatentIn Ver. 2.0
/
/ SEQ ID NO 39
/
/ LENGTH: 1443
/
/ TYPE: PRT
/
/ ORGANISM: Artificial Sequence
/
/ FEATURE:
/
/ OTHER INFORMATION: Description of Artificial Sequence: factor VIII lacking
/
/ OTHER INFORMATION: the B domain
/
/ US-10-131-510A-39

```

QY : 1 AYKVDSCEPPEQLIMKNNEEAEDYDDL 29
1 : | : | : | | | | | | : | : | | | | : |
Dh 342 AHVRVSCAEEPOLRRKADEF-EDYDDNL 369

Search completed: October 21, 2003, 19:01:22
Job time : 23.25 secs